OPTIMAL LIQUIDATION IN A LEVEL-I LIMIT ORDER BOOK FOR LARGE-TICK STOCKS

ANTOINE JACQUIER AND HAO LIU

Abstract. We propose a framework to study the optimal liquidation strategy in a limit order book for large-tick stocks, with the spread equal to one tick. All order book events (market orders, limit orders and cancellations) occur according to independent Poisson processes, with parameters depending on price move directions. Our goal is to maximise the expected terminal wealth of an agent who needs to liquidate her positions within a fixed time horizon. By assuming that the agent trades (through sell limit order or/and sell market order) only when the price moves, we model her liquidation procedure as a semi-Markov decision process, and compute the semi-Markov kernel using Laplace method in the language of queueing theory. The optimal liquidation policy is then solved by dynamic programming, and illustrated numerically.

1. Introduction

Nowadays, most equity and derivative exchanges all over the world are using order-driven trading mechanisms: Euronext, Helsinki, Hong Kong, Tokyo, Toronto are pure order-driven markets and NYSE, Nasdaq, LSE are hybrid markets [23]. Different from a quote-driven market, where large market makers centralise buy and sell orders and provide liquidity to other market participants through setting the bid and ask quotes, an order-driven market is much more flexible, which allows all market participants to send buy or sell orders specifying the price and amount they want to trade into a limit order book (LOB). According to the classical terminology [17, Section II.B], orders leading to an immediate execution upon submission based on the LOB's trade-matching algorithm are called market orders, while orders that do not result in an immediate execution and therefore are stored in the LOB are called limit orders. The active limit orders can either get executed by subsequent counterpart market orders based on a certain priority rule[1] or be cancelled. Therefore, a LOB can be understood as a collection of buy and sell limit orders stored at different price levels awaiting to be executed by counterpart market orders or cancellations.

Because of the popularity of the order-driven markets, many studies have recently been conducted on the optimal execution strategies in LOBs. In fact, theoretical modelling of the complete dynamics of a LOB is extremely complicated, and currently there is no clear consensus on a universal approach. This makes constructing an optimal trading strategy in a LOB even more difficult and one has to rely on simplifying assumptions for the

Date: February 21, 2017.

2010 *Mathematics Subject Classification.* 91G60, 91G99, 60K15, 60K20.

Key words and phrases. limit order book, optimal liquidation, semi-Markov decision process, queueing theory, dynamic programming.

The authors would like to thank Martin Gould and Fabrizio Lillo for useful discussions. AJ acknowledges financial support from the EPSRC First Grant EP/M008436/1 and HL acknowledges financial support from SHELL. .

[1] A priority rule regulates how limit orders stored at the same price level in the LOB will get executed. By far the most common priority rule is 'price-time' [17, Section II. D], that is, limit orders posted closer to the mid price will get priority and limit orders posted at the same price follow the 'first come first serve' rule.

order book, as well as setting constraints on the admissible strategies. For example, Obizhaeva and Wang [29] and Alfonsi, Fruth and Schied [4] focus on the optimal execution strategies solely using market orders, assuming continuous price levels and exponential resilience of the LOB. Avellaneda and Stoikov [6] and Guilbaud and Pham [19] construct an optimal market making policy for a small agent by assuming that the mid price of the asset is an exogenous Markov process and the agent's limit orders are matched by the market counterparts at Poisson rates. Maglaras, Moallemi and Zheng [25] formulate an optimal liquidation strategy consisting of a limit order placement at inception together with block and flow market order submissions over a short time horizon. Fodra and Pham [15] model the asset price as a Markov renewal process and derive an optimal market making strategy consisting of continuous placement of limit orders. Nevmyvaka, Feng and Kearns [28] apply reinforcement learning to search for an optimal liquidation strategy by restricting the agent's actions to repositions of all the remaining inventory as a limit order and assuming the agent's actions do not change the LOB state.

The purpose of this paper is to formulate and solve a stylised optimal liquidation problem for an agent who wants to sell a pre-specified quantity of stocks over a fixed (short) trading window in a LOB where the price-time priority mechanism is applied. Information available to this agent contains historical order flows and states of the LOB at the best prices ('Level-I' data). In particular, we are mostly interested in how different trading conditions (LOB state, inventory position, time to maturity) impact the agent's decisions. In order to achieve this, we first build up a 'Level-I' LOB model describing the trading environment whose dynamics are driven by the general market participants' order flows and exogenous information. Realistic simplifying assumptions for this LOB follow those in [11, 12], including unit order size, constant one-tick spread, Poisson order flows, depletion of the best bid (resp. ask) queue moving the price one tick downward (resp. upward) and volumes at best prices after a price move regarded as stationary variables drawn from a joint distribution. We further develop this model by allowing the Poisson rates of the order flows and the joint distribution determining the size of the volumes at the best prices after a price move to depend on the price move direction. Under these assumptions, the evolution of this LOB can be modelled as a Markov renewal process [16], whose transition mechanism is intuitively described by a queueing race between the volumes at the best prices. We then consider a risk-neutral agent who tries to maximise her expected terminal wealth by selling a fixed amount of stock within a fixed (finite) time horizon in this LOB. In order to model the price-time priority rule and capture the execution of the agent's limit orders, the agent can only submit a market order at the best bid price and/or post a limit order at the best ask price immediately after the price moves, and her market orders never consume up the volumes at the best bid price. The agent's trading procedure is then formulated through a (stationary) semi-Markov decision process within a finite horizon [21], among a certain class of deterministic stationary policies. In general, at each price-change time, the optimal policy is a deterministic function which tells the agent the size of the market and limit order to trade based on the current LOB state (price move direction, volumes at the best prices), the agent's inventory position and time to maturity in order to achieve terminal wealth maximisation.

This framework is particularly suitable for large-tick stocks (see [7, Section 4] for definition and selection criteria). Indeed, the bid-ask spread of a LOB for large-tick assets is almost always equal to one tick [13] and such a LOB is in general fully occupied around the mid price. Furthermore, based on the fact that only about 0.1% to 0.5% of market orders on large-tick stocks exceed the size of the best queue volume [8], our constraints on the agent's market order strategy are reasonable.

This paper is organised as follows. In Section 2, we set the basic assumptions and illustrate the evolutional dynamics of a 'Level-I' LOB and define the objective as well as the admissible trading strategy set for the agent. In Section 3, a semi-Markov decision process with a horizon-related deterministic stationary policy is introduced to model the agent's trading procedure and an optimal policy is defined. In Section 4, we provide an expression for the semi-Markov kernel, which works as the transition mechanism of the semi-Markov decision process. Existence of the optimal policy are proved in Section 5, and empirical studies show our numerical results in Section 6.

Notations: we shall use the following notations: $\mathbb{N} := \{0, 1, 2, \dots\}$, $\mathbb{N}^+ := \{1, 2, \dots\}$, $\mathbb{R}^+ := (0, \infty)$, $\mathbb{R}_0^+ := \mathbb{R}^+ \cup \{0\}$ and $\mathbb{R}^- := \mathbb{R} \setminus \mathbb{R}_0^+$. In this paper, $T > 0$ is a fixed terminal time, and we denote $\mathbb{T} := [0, T]$ and $\mathbb{T}_- := \mathbb{R}^- \cup \mathbb{T}$. For a continuous-time process $(L_s)_{s \geq 0}$, denote τ_L its first passage time to the origin, and f_L (resp. F_L) the density (resp. cumulative distribution function) of τ_L and, as usual $\overline{F}_L := 1 - F_L$.

2. Limit Order Book and Trading Strategy

2.1. **'Level-I' Limit Order Book Model.** We consider a limit order book characterised by two resolution parameters as in [17]: the tick size $\varepsilon > 0$ represents the smallest interval (assumed constant) between price levels, and the lot size, $\sigma > 0$, specifies the smallest amount of the asset that can be traded. All buy and sell orders thus must arrive at a price $k_1\varepsilon$ and with a size $k_2\sigma$, for some $k_1, k_2 \in \mathbb{N}^+$. Throughout this paper we shall work with the following modelling assumptions for the limit order book:

Assumption 2.1 (Order book settings)**.**

(a) orders from general market participants are of unit size, defined by σ actual size;
(b) the spread of the limit order book is equal to the tick size ε.

In particular, this limit order book is modelled on a 'Level-I' data, that is, the volumes and order flows at the best bid and ask prices. This is motivated by the fact that the imbalance between the order flows at the best bid and ask prices is shown to be a good predictor of the order book dynamics [9, 11], together with the fact that data at the best prices are more obtainable than the 'Level-II' data. Following [10], we impose the following assumptions for the evolution of the limit order book:

Assumption 2.2 (Evolution of the limit order book)**.**

(a) whenever orders at the best bid (resp. ask) price are depleted, both the best bid and ask prices decrease (resp. increase) by one tick;
(b) immediately after each price increase (resp. decrease), volumes at the best bid and ask prices are treated as random variables with joint distribution f_{+1} (resp. f_{-1}) : $(\mathbb{N}^+)^2 \to [0, 1]$; for any $x_1, x_2 \in \mathbb{N}^+$, $f_{+1}(x_1, x_2)$ (resp. $f_{-1}(x_1, x_2)$) represents the probability that the best bid and ask queue contain x_1 and x_2 unit limit orders (of actual size $x_1\sigma$ and $x_2\sigma$), right after a price increase (resp. decrease).

Remark 2.3. Assumption 2.2 presumes that the limit order book contains no empty level near mid price so that price changes are restricted to one tick, and that price changes are entirely due to exogenous information, in which case market participants swiftly readjust their order flows at the new best prices, as if a new state of the limit order book is drawn from its invariant distribution [20]. In other words, we rule out the possibility that depletion of the best bid (resp. ask) queue is followed by an insertion of buy (resp. sell) limit orders inside the spread, keeping the best bid and ask prices unchanged.

Modelling order flows from general market participants is based on the 'zero-intelligence' approach [10, 12, 31].

Assumption 2.4 (Poisson order flows)**.** All order book events (market orders, limit orders and cancellations) from general market participants occur according to independent Poisson processes, with parameters depending on the price move direction. To be more specific, taking order flows at the best ask price for example, during any period between a price increase (resp. decrease) and the next price change, the following mutually independent events happen:

(a) market buy orders arrive at independent, exponential times with rate $\mu^a_{+1} > 0$ (resp. $\mu^a_{-1} > 0$);
(b) limit sell orders arrive at independent, exponential times with rate $\kappa^a_{+1} > 0$ (resp. $\kappa^a_{-1} > 0$);
(c) cancellations of limit orders occur at independent, exponential times with rate $\theta^a_{+1} > 0$ (resp. $\theta^a_{-1} > 0$) multiplied by the amount (in unit size) of the outstanding sell limit orders.

We assume an analogous framework at the best bid price, with parameters $\mu^b_{+1}, \mu^b_{-1}, \kappa^b_{+1}, \kappa^b_{-1}, \theta^b_{+1}, \theta^b_{-1} > 0$.

Remark 2.5.

- Although the 'zero-intelligence' model is not exactly compatible with empirical observations [34], it still retains the major statistical features of limit order books while remaining computationally manageable. With the 'zero-intelligence' hypothesis, the agent can easily characterise the dynamical properties of the limit order book from historical data without assuming behavioural assumptions or resorting to unobservable parameters for other market participants.
- Assumption 2.4(c) means that if there are v limit orders at the best ask (resp. bid) price, each of which can be cancelled at an exponential time with rate θ^a (resp. θ^b) independently, and the overall cancellation rate is then $\theta^a v$ (resp. $\theta^b v$).

2.2. **Objective and admissible trading strategies.** In the limit order book model introduced in Section 2.1, we assume that the agent is risk-neutral and her goal is to maximise the expected wealth obtained through selling the stock of $\chi \in \mathbb{N}^+$ unit size ($\chi\sigma$ actual size) within the finite horizon $\mathbb{T}$. The following assumption describes the set of admissible trading strategies:

Assumption 2.6 (Admissible trading strategies)**.**

(a) the agent can only trade immediately after a price change; let τ_n denote her n-th decision epoch, namely the time of the n-th price change; $\tau_0 = 0$ and the last decision epoch before or at maturity is $\tau_{\mathfrak{n}}$, where $\mathfrak{n} := \sup\{n \in \mathbb{N} : \tau_n \le T\}$;
(b) at maturity T, the agent is required to sell all the unexecuted stocks through a market order;
(c) at each decision epoch τ_n, the agent observes the bid and ask queues, with volumes of v^b and v^a unit size; she can then post a sell limit order of l unit size at the best ask price and submit a sell market order of m unit size at the best bid price; we assume that the best bid queue is never depleted by the agent, and that the agent is slow, meaning that her limit order (of l unit size) has less time priority upon submission than the limit orders from other market participants (of v^a unit size);
(d) the agent follows a 'no cancellation' rule: she will not cancel her limit order unless the price goes down;
(e) short selling is not allowed.

Restricting the agent's trading actions at price changes (Assumption 2.6(a)) might sound relatively strong, but is necessary to capture the time-priority rule and the executions of the agent's limit orders. We shall study later in Section 3.3 how to define an optimal policy maximising the expected wealth at maturity T.

3. Trading procedure modelled by semi-Markov decision processes

A semi-Markov decision model [33, Chapter 7] is a dynamic system whose states are observed at random epochs, each of when an action is taken and a payoff incurs (either as a lump sum at that epoch or at a rate continuously until the next epoch) as a result of the action. It also satisfies the following two Markovian properties:

(M1) given the current state and the action at a given epoch, the time until the next epoch and the next state only depend on the current state and action;

(M2) the payoff incurred at any epoch depends only on the state and the action at that epoch.

The semi-Markov decision model well describes the agent's liquidation problem within our stylised limit order book: the limit order book with the agent's participation is a dynamic system, and the agent's selling action at each decision epoch may lead to a payoff. Indeed, Assumption 2.6(a) enables us to track the state of this system merely at the decision epochs, and Assumptions 2.2, 2.4 and 2.6(c) ensure that the transition mechanism of the system is stationary and satisfies **(M1)**-**(M2)**. Moreover, according to Assumption 3.3, each payoff from the agent's matched limit order is allocated to the nearest incoming decision epoch in order to make the payoff as a lump sum. In Section 3.1, we define a (stationary) semi-Markov decision model with lump-sum payoffs for the agent's liquidation procedure. In Section 3.2, we define a horizon-related deterministic stationary policy and illustrate the evolution of the semi-Markov decision process. In Section 3.3, we give the definition of the expected reward function, the value function and the optimal policy for the agent's liquidation problem.

3.1. Semi-Markov decision model.

The semi-Markov decision model with lump-sum payoffs and the finite-horizon constraint is defined as a six-tuple $\{\mathcal{E}, (\mathcal{A}(e))_{e\in\mathcal{E}}, Q(\cdot,\cdot|\cdot), P(\cdot|\cdot), r(\cdot,\cdot), w(\cdot,\cdot)\}$, where each element is defined below.

3.1.1. *State space.* Fix $N \in \mathbb{N}^+$ large enough. The state space $\mathcal{E} := \{-1,+1\} \times \{1,\dots,N^+\}^3 \times \{0,\dots,N\}^2$ is the set of all pre-decision conditions of the system (i.e. the limit order book with the agent's participation) observed at each decision epoch. Specifically, the system being in state $e := (j, v^b, v^a, p, z, y) \in \mathcal{E}$ means that:

- the ask/bid price change is equal to j tick;
- the best bid (resp. ask) queue contains v^b (resp. v^a) unit orders;
- the ask price[2] is equal to $p\varepsilon$;
- the executed part of the limit order posted by the agent at the previous decision epoch is of z unit size;
- the agent's remaining inventory position is of y unit size.

3.1.2. *Action space.* The action space $\mathcal{A} := \{0,\dots,\overline{m}\} \times \{0,\dots,\overline{l}\}$, with $\overline{m}, \overline{l} \in \mathbb{N}^+$, represents the set of trading strategies, that is, the amount (in unit size) of the market and limit order that the agent chooses to submit and post at the best bid and ask price respectively. The constant $\overline{m}$ (resp. $\overline{l}$) represents the maximum amount (in unit size) of a single market (resp. limit) order that the agent is allowed to trade. From Assumption 2.6(c)(e), the agent's admissible action space in state $e \in \mathcal{E}$ is defined by

$$\mathcal{A}(e) := \left\{(m,l) \in \mathcal{A} : m < v^b, m + l \le y\right\},$$

[2] The stylised limit order book model doesn't implement a positive restriction on the stock price. But we assume that the stock price is far above zero at inception and the liquidation horizon $\mathbb{T}$ is short, so that the stock price will never become negative.

so that the agent will never consume up the entire best bid queue nor short sell. The set of all feasible state-action pairs is denoted by $\mathcal{K} := \{(e,\alpha)|e \in \mathcal{E}, \alpha \in \mathcal{A}(e)\}$.

3.1.3. *Semi-Markov kernel.* Before introducing our next concept, recall the following definition.

Definition 3.1 (sub-/semi-Markov kernel). Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be real measurable spaces. A map $p(\cdot|\cdot) : \mathcal{F}_2 \times \Omega_1 \to [0,1]$ is called a sub-Markov kernel on Ω_2 given Ω_1 if:

- for any $\omega_1 \in \Omega_1$, $p(\cdot|\omega_1)$ is a measure on $(\Omega_2, \mathcal{F}_2)$ with $p(\Omega_2|\omega_1) \le 1$;
- for any $F_2 \in \mathcal{F}_2$, $p(F_2|\cdot)$ is a Borel measurable function.

In particular, if $p(\Omega_2|\omega_1) = 1$ for all $\omega_1 \in \Omega_1$, then $p(\cdot|\cdot)$ is a Markov kernel on Ω_2 given Ω_1. Furthermore, a map $q(\cdot,\cdot|\cdot) : \mathbb{R}_0^+ \times \mathcal{F}_2 \times \Omega_1 \to [0,1]$ is a semi-Markov kernel on $\mathbb{R}_0^+ \times \Omega_2$ given Ω_1 if:

- for $(F_2, \omega_1) \in \mathcal{F}_2 \times \Omega_1$, $q(\cdot, F_2|\omega_1)$ is non-decreasing, right-continuous and $q(0, F_2|\omega_1) = 0$;
- for $t \ge 0$, $q(t,\cdot|\cdot)$ is a sub-Markov kernel on Ω_2 given Ω_1;
- the limit $\lim_{t\uparrow\infty} q(t,\cdot|\cdot)$ is a Markov kernel on Ω_2 given Ω_1.

In our model, let $Q(\cdot,\cdot|\cdot)$ be a semi-Markov kernel on $\mathbb{R}_0^+ \times \mathcal{E}$ given $\mathcal{K}$, determining the (stationary) transition mechanism of the semi-Markov decision process: for any $t \ge 0$ and $\tilde{e} \in \mathcal{E}$, given the state-action pair $(e,\alpha) \in \mathcal{K}$ at some decision epoch, the quantity[3] $Q(t, \tilde{e}|(e,\alpha))$ represents the (joint) probability that the time until the next decision epoch is less than or equal to t and the next system state is $\tilde{e}$. Detailed computations are given in Section 4.

3.1.4. *Terminal kernel.* The terminal kernel $P(\cdot|\cdot)$ is a sub-Markov kernel on $\mathbb{N}$ given $\mathcal{K} \times \mathbb{T}_-$, and describes the execution dynamics between the last decision epoch and the maturity: for any $\mathfrak{z} \in \mathbb{N}$, given the state-action pair $(e,\alpha) \in \mathcal{K}$ and the time to maturity $\lambda \in \mathbb{T}_-$ at some decision epoch[4], the quantity[5] $P(\mathfrak{z}|(e,\alpha),\lambda)$ represents the (joint) probability that the time until the next decision epoch is strictly larger than λ and the executed part of the limit order up to the maturity is of $\mathfrak{z}$ unit size. Detailed computations are given in Section 4.

Remark 3.2. According to our modelling framework, the terminal kernel satisfies the following properties:

- $P(0|(e,\alpha),\lambda) = 1$ when $\lambda \le 0$;
- $\sum_{\mathfrak{z}\ge 0} P(\mathfrak{z}|(e,\alpha),\lambda) = 1 - Q(\lambda, \mathcal{E}|(e,\alpha))$ when $\lambda > 0$;
- $P(\mathfrak{z}|(e,\alpha),\lambda) = 0$ when $\mathfrak{z} > l$;

for any $(e,\alpha) \in \mathcal{K}$.

3.1.5. *Periodical reward function.* The periodical reward function $r : \mathcal{K} \to \mathbb{R}_0^+$ is defined as

$$r(e,\alpha) := \rho\left[m\,(p-1) + z\,(p-j)\right], \qquad \text{for all } (e,\alpha) \in \mathcal{K}, \qquad \text{where } \rho := \varepsilon\sigma, \tag{3.1}$$

and represents the lump-sum payoff associated with a decision epoch given the state-action pair (e,α). Specifically, the definition (3.1) is given based on the following assumption that assigns the payoff from the matched part of the agent's limit order to the nearest incoming decision epoch.

Assumption 3.3 (Periodic reward function). For $n \in \mathbb{N}^+$, the payoff from the matched limit order within the interval $[\tau_{n-1}, \tau_n)$ is allocated at τ_n.

[3]By abuse of language, we write $Q(t, \{\tilde{e}\}|(e,\alpha))$ as $Q(t, \tilde{e}|(e,\alpha))$.

[4]A decision epoch with time to maturity $\lambda < 0$ means that it happens a period of time $|\lambda|$ after the maturity.

[5]By abuse of language, we write $P(\{\mathfrak{z}\}|((e,\alpha),\lambda))$ as $P(\mathfrak{z}|(e,\alpha),\lambda)$.

Assuming that the system is in state $e \in \mathcal{E}$ and the agent takes action $\alpha \in \mathcal{A}(e)$ at some decision epoch. She then earns an immediate payoff worth $m(p-1)\rho$ from submitting the market order of m unit size at the best bid price $(p-1)\varepsilon$. On top of that, the matched limit order of z unit size at the previous best ask price $(p-j)\varepsilon$ entails a payoff worth $z(p-j)\rho$, which is allocated at the current decision epoch according to Assumption 3.3.

3.1.6. *Terminal reward function.* The terminal reward function $w : \mathcal{K} \times \mathbb{N} \to \mathbb{R}_0^+$ is defined as

$$w(e, \alpha, \mathfrak{z}) := \rho\left[(p-1)(y-m) + \mathfrak{z}\right] - g(y - m - \mathfrak{z}), \qquad \text{for all } (e,\alpha) \in \mathcal{K} \text{ and } \mathfrak{z} \in \mathbb{N}, \tag{3.2}$$

where the market impact function $g : \mathbb{N} \to \mathbb{R}_0^+$ is of the form

$$g(x) := \rho \frac{x}{\overline{v}}, \tag{3.3}$$

for a constant $\overline{v} \in \mathbb{N}^+$. For any $(e, \alpha) \in \mathcal{K}$ and $\mathfrak{z} \in \mathbb{N}$, the quantity $w(e, \alpha, \mathfrak{z})$ represents the lump-sum payoff associated with the maturity T, given the state-action pair (e, α) at the last decision epoch, and the matched part of the agent's limit order between the last decision epoch and the maturity being of $\mathfrak{z}$ unit size. Particularly, the identity (3.2) is given based on the following assumption:

Assumption 3.4 (Terminal reward function)**.**

(a) the payoff from the matched limit order obtained within the interval $[\tau_{\mathfrak{n}}, T)$ is allocated at T;
(b) when depicting the market impact brought by the market order at maturity, we assume that the impact is linear with $\overline{v}$ representing the average depth (in unit size) on the bid side of the limit order book;
(c) the unexecuted shares at maturity cannot sweep all the liquidity on the bid side of the limit order book, so that the terminal reward function is $\mathbb{R}_0^+$-valued.

Assumption 3.4(b) yields the market impact function g in Formula (3.3). Furthermore, based on Assumption 3.4(a)(b), the terminal reward $w(e, \alpha, \mathfrak{z})$ consists of the payoff from the matched limit order (of amount $\rho p \mathfrak{z}$) and the market order at maturity (of amount $\rho(p-1)(y-m-\mathfrak{z})$), deducted by the corresponding market impact (of amount $g(y - m - \mathfrak{z})$).

3.2. **Dynamics of the finite-horizon semi-Markov decision process.** We assume that the agent applies a horizon-related deterministic stationary policy defined below, specifying a decision rule for her actions.

Definition 3.5. A horizon-related deterministic stationary policy is a measurable function

$$\pi : \mathcal{E} \times \mathbb{T}_{-} \ni (e, \lambda) \mapsto \alpha \in \mathcal{A}(e),$$

such that $\pi(e, \lambda) = (0, 0)$ for any $(e, \lambda) \in \mathcal{E} \times \mathbb{R}^-$. We further denote by Π the set of all horizon-related deterministic stationary policies.

Remark 3.6. At the n-th decision epoch, with system state $e_n \in \mathcal{E}$ and time to maturity $\lambda_n := T - \tau_n$, an action $\pi(e_n, \lambda_n)$ is given by the policy π. In particular, at any decision epoch τ_n with $n > \mathfrak{n}$ (namely $\lambda_n < 0$), the agent stops trading as $\alpha_n = (0, 0)$, following Assumption 2.6(b).

Table 1 summarises the evolution of the semi-Markov decision model with a deterministic stationary policy π. Suppose that the system is in state e_0 at τ_0, and the agent has a planned trading horizon λ_0. According to the policy π, she chooses an action $\alpha_0 = \pi(e_0, \lambda_0)$. It then takes a period of time t_1 to reach the next decision epoch ($\tau_1 = \tau_0 + t_1$), at which point the system state changes to e_1 and the time to maturity for the agent becomes $\lambda_1 = \lambda_0 - t_1$. She then chooses the action $\alpha_1 = \pi(e_1, \lambda_1)$, and so on. At the n-th decision epoch, a periodic

Index	Time	State	Time to Maturity	Action	Payoff
Initial	τ_0	e_0	$\lambda_0 \geq 0$	$\alpha_0 = \pi(e_0, \lambda_0)$	$r(e_0, \alpha_0)$
1st	$\tau_1 = \tau_0 + t_1$	e_1	$\lambda_1 = \lambda_0 - t_1 \geq 0$	$\alpha_1 = \pi(e_1, \lambda_1)$	$r(e_1, \alpha_1)$
2nd	$\tau_2 = \tau_1 + t_2$	e_2	$\lambda_2 = \lambda_1 - t_2 \geq 0$	$\alpha_2 = \pi(e_2, \lambda_2)$	$r(e_2, \alpha_2)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(\mathfrak{n}-1)$-th	$\tau_{\mathfrak{n}-1} = \tau_{\mathfrak{n}-2} + t_{\mathfrak{n}-1}$	$e_{\mathfrak{n}-1}$	$\lambda_{\mathfrak{n}-1} = \lambda_{\mathfrak{n}-2} - t_{\mathfrak{n}-1} \geq 0$	$\alpha_{\mathfrak{n}-1} = \pi(e_{\mathfrak{n}-1}, \lambda_{\mathfrak{n}-1})$	$r(e_{\mathfrak{n}-1}, \alpha_{\mathfrak{n}-1})$
$\mathfrak{n}$-th	$\tau_{\mathfrak{n}} = \tau_{\mathfrak{n}-1} + t_{\mathfrak{n}}$	$e_{\mathfrak{n}}$	$\lambda_{\mathfrak{n}} = \lambda_{\mathfrak{n}-1} - t_{\mathfrak{n}} \geq 0$	$\alpha_{\mathfrak{n}} = \pi(e_{\mathfrak{n}}, \lambda_{\mathfrak{n}})$	$r(e_{\mathfrak{n}}, \alpha_{\mathfrak{n}})$
	Terminal T				$w(e_{\mathfrak{n}}, \alpha_{\mathfrak{n}}, \mathfrak{z})$
$(\mathfrak{n}+1)$-th	$\tau_{\mathfrak{n}+1} = \tau_{\mathfrak{n}} + t_{\mathfrak{n}+1}$	$e_{\mathfrak{n}+1}$	$\lambda_{\mathfrak{n}+1} = \lambda_{\mathfrak{n}} - t_{\mathfrak{n}+1} < 0$	$\alpha_{\mathfrak{n}+1} = (0,0)$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

TABLE 1. Evolution of the semi-Markov decision process

payoff of amount $r(e_n, \alpha_n)$ incurs. At maturity T, a terminal payoff $w(e_{\mathfrak{n}}, \alpha_{\mathfrak{n}}, \mathfrak{z})$ is obtained. In particular, after T ($n > \mathfrak{n}$), the agent takes no action ($\alpha_n = (0,0)$), and no payoff is paid. In the following, we construct the semi-Markov decision process in a probability space based on the Ionescu Tulcea's Theorem.

Definition 3.7. Let $(\Omega, \mathcal{F})$ be a measurable space consisting of the sample space Ω, defined by

$$\Omega := \left\{\mathfrak{n} \in \mathbb{N}, \mathfrak{z} \in \mathbb{N}, \left(\{t_n, e_n, \lambda_n, \alpha_n\} \in \mathbb{R}_0^+ \times \mathcal{E} \times \mathbb{T}_- \times \mathcal{A}(e_n)\right)_{n\in\mathbb{N}}\right\},$$

and the corresponding Borel σ-algebra $\mathcal{F}$. Define the random variables $\mathfrak{N}$, $\mathfrak{Z}$, X_n, E_n, Λ_n, A_n on $(\Omega, \mathcal{F})$ as:

$$\begin{aligned} \mathfrak{N}(\omega) &= \mathfrak{n}, & \mathfrak{Z}(\omega) &= \mathfrak{z}, \\ X_n(\omega) &= t_n, & E_n(\omega) &:= \left(J_n, V_n^b, V_n^a, P_n, Z_n, Y_n\right)(\omega) = e_n, \\ \Lambda_n(\omega) &= \lambda_n, & A_n(\omega) &:= (M_n, L_n)(\omega) = \alpha_n, \end{aligned}$$

for any $\omega \in \Omega$ and $n \in \mathbb{N}$, where

- X_n is the time between the $(n-1)$-th and the n-th decision epoch ($X_0 = 0$ almost surely);
- E_n, Λ_n, A_n represent the system state, time to maturity and agent's action at the n-th decision epoch;
- $\mathfrak{N}$ is the index of the last decision epoch;
- $\mathfrak{Z}$ is the amount (in unit size) of the agent's limit order executed between the $\mathfrak{N}$-th decision epoch and the maturity.

Remark 3.8. Based on this modelling framework, the following properties hold almost surely for $n \in \mathbb{N}$

- $\Lambda_{n+1} = \Lambda_n - X_{n+1}$: evolution of the time to maturity;
- $P_{n+1} = P_n + J_{n+1}$: evolution of the ask price (in tick size);
- $Y_{n+1} = Y_n - M_n - Z_{n+1}$: evolution of the inventory position (in unit size);
- $Z_{n+1} \leq L_n$: the amount of the matched limit order cannot exceed that of the limit order posted by the agent in each queueing race;
- $\mathfrak{N} = \sup\{n \in \mathbb{N} : \Lambda_n \geq 0\}$: index of the last decision epoch;
- $\mathfrak{Z} \leq Z_{\mathfrak{N}+1}$: the amount of the matched limit order between the last decision epoch and the maturity cannot exceed that of limit order executed when there is no finite-horizon restriction.

Theorem 3.9. *[Tulcea's Theorem* [5, Section 2.7.2]*] For any $(e,\lambda) \in \mathcal{E} \times \mathbb{T}$ and $\pi \in \Pi$, there exists a unique probability measure $\mathbb{P}^{\pi}_{(e,\lambda)}$ on $(\Omega, \mathcal{F})$ such that, for any $t \geq 0$, $\tilde{e} \in \mathcal{E}$, $\alpha \in \mathcal{A}$, $\mathfrak{z} \in \mathbb{N}$ and $n \in \mathbb{N}$,*

$$\begin{aligned} \mathbb{P}^{\pi}_{(e,\lambda)}(X_0 = 0, E_0 = e, \Lambda_0 = \lambda) &= 1, \\ \mathbb{P}^{\pi}_{(e,\lambda)}(A_n = \alpha | H_n = h_n) &= \mathbb{1}_{\{\pi(e_n, \lambda_n) = \alpha\}}, \\ \mathbb{P}^{\pi}_{(e,\lambda)}(X_{n+1} \leq t, E_{n+1} = \tilde{e} | H_n = h_n, A_n = \alpha_n) &= Q(t, \tilde{e} | (e_n, \alpha_n)), \\ \mathbb{P}^{\pi}_{(e,\lambda)}(X_{n+1} > \lambda_n, \mathfrak{Z} = \mathfrak{z} | H_n = h_n, A_n = \alpha_n) &= P(\mathfrak{z} | (e_n, \alpha_n), \lambda_n), \end{aligned}$$

where

$$H_n := \begin{cases} (\{X_0, E_0, \Lambda_0\}), & \text{if } n = 0, \\ \big(\{X_i, E_i, \Lambda_i, A_i\}_{i=0,\dots,n-1}, \{X_n, E_n, \Lambda_n\}\big), & \text{if } n \in \mathbb{N}^+, \end{cases}$$

is the sequence of random variables describing the history up to the n-th decision epoch (realisations of the random variables (or sequences of random variables) are denoted by the corresponding lower case letters).

3.3. **Value function and optimal policy.** Consider an agent with objective and trading strategies as in Section 2.2, introduce the following definition.

Definition 3.10. Define the finite-horizon expected reward function under a policy $\pi \in \Pi$ by

$$V^{\pi}(e,\lambda) := \mathbb{E}^{\pi}_{(e,\lambda)} \left(\sum_{n=0}^{\mathfrak{N}} r(E_n, A_n) + w(E_{\mathfrak{N}}, A_{\mathfrak{N}}, \mathfrak{Z}) \right), \qquad \text{for any } (e,\lambda) \in \mathcal{E} \times \mathbb{T}, \tag{3.4}$$

as well as the value function

$$V^{*}(e,\lambda) := \sup \{ V^{\pi}(e,\lambda), \pi \in \Pi \}. \tag{3.5}$$

A policy $\pi^* \in \Pi$ is called $\mathbb{T}$-optimal if the equality

$$V^{\pi^*}(e,\lambda) = V^{*}(e,\lambda) \tag{3.6}$$

holds for all $(e,\lambda) \in \mathcal{E} \times \mathbb{T}$.

Remark 3.11. For any $(e,\lambda) \in \mathcal{E} \times \mathbb{T}$, we can rewrite the quantity $V^{\pi}(e,\lambda)$ in (3.4) as

$$\begin{aligned} V^{\pi}(e,\lambda) &= \mathbb{E}^{\pi}_{(e,\lambda)} \left(\sum_{n=0}^{\infty} \big(r(E_n, A_n) \mathbb{1}_{\{\mathfrak{N} \geq n\}} \big) + w(E_n, A_n, \mathfrak{Z}) \mathbb{1}_{\{\mathfrak{N} = n\}} \right) \\ &= \mathbb{E}^{\pi}_{(e,\lambda)} \left(\sum_{n=0}^{\infty} \big(r(E_n, A_n) \mathbb{1}_{\{\Lambda_n \geq 0\}} + w(E_n, A_n, \mathfrak{Z}) \mathbb{1}_{\{0 \leq \Lambda_n < X_{n+1}\}} \big) \right) \\ &= \sum_{n=0}^{\infty} \mathbb{E}^{\pi}_{(e,\lambda)} \Big(r(E_n, A_n) \mathbb{1}_{\{\Lambda_n \geq 0\}} + w(E_n, A_n, \mathfrak{Z}) \mathbb{1}_{\{0 \leq \Lambda_n < X_{n+1}\}} \Big), \end{aligned} \tag{3.7}$$

where the second equality follows by writing

$$\begin{aligned} &\{\mathfrak{N} \geq n\} = \{\Lambda_0 \geq 0, \dots, \Lambda_n \geq 0\} = \{\Lambda_n \geq 0\}, \\ &\{\mathfrak{N} = n\} = \{\Lambda_0 \geq 0, \dots, \Lambda_n \geq 0, \Lambda_{n+1} < 0\} = \{\Lambda_n \geq 0, \Lambda_{n+1} < 0\} = \{0 \leq \Lambda_n < X_{n+1}\}, \end{aligned}$$

since the sequence $\{\Lambda_n\}_{n \in \mathbb{N}}$ is non-increasing, and the third equality is due to the non-negativity of the periodic/terminal reward function and the monotone convergence theorem.

4. Semi-Markov kernel

We now provide the expressions for the semi-Markov kernel $Q(\cdot,\cdot|\cdot)$ and the terminal kernel $P(\cdot|\cdot)$ defined in Section 3.1 using the language of queueing theory. We first (Section 4.1) model the dynamics of the best queues with the agent's participation as generalised birth-death processes, and derive the closed-form expressions for the semi-Markov kernel and the terminal kernel in all possible scenarios in terms of the distributions of the first-passage time of the generalised birth-death processes to zero. We then (Section 4.2) compute these distributions by using Laplace method.

4.1. **Closed-form expressions.** For notational convenience, we shall fix an element $(\mathfrak{e},\lambda)$ in $\mathcal{E}\times\mathbb{T}$ together with a deterministic stationary policy $\pi\in\Pi$ and denote $\mathbb{P}^{\pi}_{(\mathfrak{e},\lambda)}$ by $\mathbb{P}$ throughout this section.

4.1.1. *Semi-Markov kernel.* According to Theorem 3.9 and the Markovian property **(M1)**, we can express the semi-Markov kernel as a (stationary) distribution of the duration and outcome of a queueing race given its initial condition and the agent's action:

$$Q(t,\tilde{e}|(e,\alpha)) = \mathbb{P}(X_{n+1}\le t, E_{n+1}=\tilde{e}|E_n=e,A_n=\alpha), \qquad \text{for any } t\ge 0, \tilde{e}\in\mathcal{E},(e,\alpha)\in\mathcal{K}, n\in\mathbb{N}. \tag{4.1}$$

To simplify further calculations, we now factorise the conditional probability in (4.1).

Proposition 4.1. *For any $\tilde{e}=\left(\tilde{j},\tilde{v}^b,\tilde{v}^a,\tilde{p},\tilde{z},\tilde{y}\right)\in\mathcal{E}$ and $\left(e:=(j,v^b,v^a,p,z,y),\alpha:=(m,l)\right)\in\mathcal{K}$, we have*

$$Q(t,\tilde{e}|(e,\alpha)) = \mathcal{Q}_{j,v,\alpha}\left(t,\tilde{j},\tilde{z}\right) f_{\tilde{j}}\left(\tilde{v}^b,\tilde{v}^a\right)\mathbb{1}_{\{\tilde{p}=p+\tilde{j}\}}\mathbb{1}_{\{\tilde{y}=y-m-\tilde{z}\}}, \tag{4.2}$$

for all $t\ge 0$, where[6] *for any $n\in\mathbb{N}$,*

$$\mathcal{Q}_{j,v,\alpha}\left(t,\tilde{j},\tilde{z}\right) := \mathbb{P}\left(X_{n+1}\le t, J_{n+1}=\tilde{j}, Z_{n+1}=\tilde{z}\middle| J_n=j,(V_n^b,V_n^a)=(v^b,v^a),A_n=\alpha\right).$$

Proof. According to Assumption 2.2 and Remark 3.8, we can write

$$\begin{aligned}
Q(t,\tilde{e}|(e,\alpha)) &= \mathbb{P}\left(X_{n+1}\le t, J_{n+1}=\tilde{j}, Z_{n+1}=\tilde{z}|E_n=e,A_n=\alpha\right)\times\\
&\qquad \mathbb{P}\left((V^b_{n+1},V^a_{n+1})=(\tilde{v}^b,\tilde{v}^a),P_{n+1}=\tilde{p},Y_{n+1}=\tilde{y}\middle|X_{n+1}\le t,J_{n+1}=\tilde{j},Z_{n+1}=\tilde{z},E_n=e,A_n=\alpha\right)\\
&= \mathcal{Q}_{j,v,\alpha}\left(t,\tilde{j},\tilde{z}\right)\mathbb{P}\left((V^b_{n+1},V^a_{n+1})=(\tilde{v}^b,\tilde{v}^a)|J_{n+1}=\tilde{j}\right)\times\\
&\qquad \mathbb{P}\left(P_{n+1}=\tilde{p}|J_{n+1}=\tilde{j},P_n=p\right)\mathbb{P}\left(Y_{n+1}=\tilde{y}|Y_n=y,M_n=m,Z_{n+1}=\tilde{z}\right)\\
&= \mathcal{Q}_{j,v,\alpha}\left(t,\tilde{j},\tilde{z}\right) f_{\tilde{j}}\left(\tilde{v}^b,\tilde{v}^a\right)\mathbb{1}_{\{\tilde{p}=p+\tilde{j}\}}\mathbb{1}_{\{\tilde{y}=y-m-\tilde{z}\}}.
\end{aligned}$$

□

Remark 4.2. The function $\mathcal{Q}$ is a semi-Markov kernel on $\mathbb{R}_0^+\times\mathcal{E}'$ given $\mathcal{K}'$, where

$$\begin{aligned}
\mathcal{E}' &:= \{-1,+1\}\times\{0,1,\dots,N\};\\
\mathcal{K}' &:= \left\{(j,v^b,v^a,\alpha): j\in\{+1,-1\},(v^b,v^a)\in\{1,\dots,N\}^2,\alpha\in\mathcal{A},m<v^b\right\}.
\end{aligned}$$

Indeed, for any $(j,v^b,v^a,\alpha)\in\mathcal{K}'$, the probability $\mathcal{Q}_{j,v,\alpha}\left(t,\{+1,-1\},\{0,\dots,l\}\right)$ converges to 1 for large t, indicating the amount of the matched limit order cannot exceed that of the limit order posted by the agent.

[6] $\mathbb{P}$ (short for $\mathbb{P}^{\pi}_{(\mathfrak{e},\lambda)}$ in this section) is the probability measure introduced in Theorem 3.9, and we use the short-hand notation $\mathbb{P}\left(X_{n+1}\le t,J_{n+1}=\tilde{j},Z_{n+1}=\tilde{z},\middle|\cdots\right)=\mathbb{P}\left(X_{n+1}\le t,J_{n+1}=\tilde{j},Z_{n+1}=\tilde{z},V^b_{n+1}\in\mathbb{N}^+,V^a_{n+1}\in\mathbb{N}^+,P_{n+1}\in\mathbb{N}^+,Y_{n+1}\in\mathbb{N}\middle|\cdots\right)$

According to Assumptions 2.2, 2.4 and 2.6, the semi-Markov kernel $\mathcal{Q}$ describes the dynamical mechanism of a queueing race between the volumes sitting at the best bid and ask prices. Intuitively, fix $\big(j, v^b, v^a, \alpha\big) \in \mathcal{K}'$, and consider a queueing race starting with v^b and v^a units limit orders (from the general market participants) at the best bid and ask prices at a certain decision epoch. The agent subsequently submits a sell market order of m unit size, which decreases the best bid volume to $(v^b - m)$ unit size, and posts a sell limit order of l unit size, which has less time priority than the pre-existing v^a units limit orders at the best ask price. After the agent's action, mutually independent order book events happen at exponential times with the rates depending on the price move direction j and therefore change the volumes of the best bid and ask queues. The queueing race terminates whenever the volume of either the best bid or ask queue reaches zero, and we denote the result of a queueing race by $+1$ (resp. -1) if the best ask (resp. bid) queue is depleted first. For $(t, \tilde{j}, \tilde{z}) \in \mathbb{R}_0^+ \times \mathcal{E}'$, the quantity $\mathcal{Q}_{j,v,\alpha}(t, \tilde{j}, \tilde{z})$ is the probability that the duration of the race is less than or equal to t, the result is $\tilde{j}$, and $\tilde{z}$ unit size of the agent's limit order gets executed. In the following, we model the dynamics of the volumes at the best bid and ask prices as generalised birth-death processes, and therefore build a connection between the semi-Markov kernel and the queueing theory.

Definition 4.3. Let $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$ be a new filtered probability space. For $v \in \mathbb{N}^+$, $l \in \mathbb{N}$ and $\kappa, \mu, \theta, \eta > 0$, define the following processes on this space:

- $(B[v, \kappa, \mu, \theta]_s)_{s \geq 0}$ is a birth and death process with state space $\mathbb{N}$ and absorbing state 0, given the initial state v; κ is the birth rate and $\mu + i\theta$ the death rate when in state $i \in \mathbb{N}^+$;
- $(C[v, l, \mu, \theta]_s)_{s \geq 0}$ is a pure death process with state space $\mathbb{N}$ and absorbing state 0 given initial state $l + v$; the death rate equals to $\mu + \max(0, i - l)\theta$ when in state $i \in \mathbb{N}^+$;
- $(G[\kappa, \mu, \theta, \eta]_s)_{s \geq 0}$ is a process with state space $\mathbb{N}$ given initial state 0. Strictly before time η, it is a birth and death process with birth rate κ and death rate $i\theta$ when in state $i \in \mathbb{N}$. After η, the birth and death rate of this process change to κ and $\mu + i\theta$ when in state $i \in \mathbb{N}^+$ and 0 becomes the absorbing state.
- $(A[v, l, \kappa, \mu, \theta]_s)_{s \geq 0}$ is a process with state space $\mathbb{N}^2$ defined by

$$A[v, l, \kappa, \mu, \theta]_s := \big(C[v, l, \mu, \theta]_s, G[\kappa, \mu, \theta, \tau_{C[v,l,\mu,\theta]}]_s\big), \quad \text{for } s \geq 0.$$

Lemma 4.4. [12, Lemma 2] *Fix $\big(j, v^b, v^a, \alpha\big) \in \mathcal{K}'$. Suppose that, at the n-th decision epoch, the queueing race starts with v^b and v^a units limit orders at the best bid and ask prices after the price moves by j tick, and the agent takes an action $\alpha = (m, l)$. On $[\tau_n, \tau_{n+1})$, define the following processes:*

- *$\widetilde{B}$: size of the orders sitting at the best bid price;*
- *$\widetilde{C}$: size of the agent's limit order together with the orders with higher time priority at the best ask price;*
- *$\widetilde{G}$: size of the orders with lower time priority than the agent's limit order at the best ask price.*

Then there exist two independent processes $B[v^b - m, \kappa_j^b, \mu_j^b, \theta_j^b]$ and $A[v^a, l, \kappa_j^a, \mu_j^a, \theta_j^a]$ such that

$$B[v^b - m, \kappa_j^b, \mu_j^b, \theta_j^b]_s = \widetilde{B}_{s+\tau_n} \qquad \text{and} \qquad A[v^a, l, \kappa_j^a, \mu_j^a, \theta_j^a]_s = (\widetilde{C}_{s+\tau_n}, \widetilde{G}_{s+\tau_n}), \qquad \text{for all } s \in [0, \tau_{n+1} - \tau_n).$$

According to Lemma 4.4, we now provide an expression for $\mathcal{Q}$, and defer its proof to Appendix A.

Proposition 4.5. *Fix $(j, v^b, v^a, \alpha) \in \mathcal{K}'$, introduce the short-hand notations:*

$$\begin{aligned} B^b &:= B[v^b - m, \kappa_j^b, \mu_j^b, \theta_j^b], & B^a &:= B[v^a, \kappa_j^a, \mu_j^a, \theta_j^a], \\ A^l &:= A[v^a, l, \kappa_j^a, \mu_j^a, \theta_j^a], & C^l &:= C[v^a, l, \mu_j^a, \theta_j^a], \end{aligned}$$

as well as the scenarios:

S1	*S2*±	*S3*	*S4*	*S5*	*S6*
$l \geq 1$ $\tilde{j} = +1$	$l = 0$ $\tilde{j} = \pm 1$	$l \geq 1$ $\tilde{j} = -1$ $\tilde{z} = 0$	$l = 1$ $\tilde{j} = -1$ $\tilde{z} = 1$	$l > 1$ $\tilde{j} = -1$ $\tilde{z} \in \{1, \dots, l-1\}$	$l > 1$ $\tilde{j} = -1$ $\tilde{z} = l$

Then the following holds for any $(t, \tilde{j}, \tilde{z}) \in \mathbb{R}_0^+ \times \mathcal{E}'$*:*

$$\mathcal{Q}_{j,v,\alpha}\left(t, \tilde{j}, \tilde{z}\right) = \begin{cases} \left[F_{A^l}(t) - \displaystyle\int_0^t f_{A^l}(u) F_{B^b}(u) \mathrm{d}u\right] \mathbb{1}_{\{\tilde{z}=l\}}, & [S1], \\ \left[F_{B^a}(t) - \displaystyle\int_0^t f_{B^a}(u) F_{B^b}(u) \mathrm{d}u\right] \mathbb{1}_{\{\tilde{z}=0\}}, & [S2+], \\ \left[F_{B^b}(t) - \displaystyle\int_0^t f_{B^b}(u) F_{B^a}(u) \mathrm{d}u\right] \mathbb{1}_{\{\tilde{z}=0\}}, & [S2-], \\ F_{B^b}(t) - \displaystyle\int_0^t f_{B^b}(u) F_{C^1}(u) \mathrm{d}u, & [S3], \\ \displaystyle\int_0^t f_{B^b}(u) \left[F_{C^1}(u) - F_{A^1}(u)\right] \mathrm{d}u, & [S4], \\ \displaystyle\int_0^t f^*_{B^b}(\epsilon) \int_0^\epsilon f^*_{C^{\tilde{z}}}(u) \mathrm{d}u \, \mathrm{d}\epsilon, & [S5], \\ \displaystyle\int_0^t f_{B^b}(u) \left[F_{C^1}(u) - F_{A^l}(u)\right] \mathrm{d}u - \sum_{z=1}^{l-1} \int_0^t f^*_{B^b}(\epsilon) \int_0^\epsilon f^*_{C^z}(u) \mathrm{d}u \, \mathrm{d}\epsilon, & [S6], \\ 0, & \textit{otherwise}, \end{cases}$$

where $f^*_{C^z}(\xi) := \mathrm{e}^{\mu_j^a \xi} f_{C^z}(\xi)$ *and* $f^*_{B^b}(\xi) := \mathrm{e}^{-\mu_j^a \xi} f_{B^b}(\xi)$ *for any* $\xi \geq 0$ *and* $z \in \mathbb{N}^+$.

4.1.2. *Terminal kernel.* According to Theorem 3.9 and the Markovian property **(M1)**, we can express the terminal kernel as

$$P(\mathfrak{z}|(e, \alpha), \lambda) = \mathbb{P}(X_{n+1} > \lambda, \mathfrak{Z} = \mathfrak{z} | E_n = e, A_n = \alpha), \tag{4.3}$$

for any $(e, \alpha) \in \mathcal{K}$, $\lambda \in \mathbb{R}$, $\mathfrak{z} \in \mathbb{N}$. Remark 3.2 implies that only the cases when $\lambda > 0$ and $\mathfrak{z} \in \{0, \dots, l\}$ need to be considered. According to Lemma 4.4, we now provide an expression for $\mathcal{Q}$, proved in Appendix B.

Proposition 4.6. *For any* $\lambda > 0, (e, \alpha) \in \mathcal{K}$ *(with the corresponding* $(j, v^b, v^a, \alpha) \in \mathcal{K}'$*)), introduce the processes* B^b, B^a, A^l, C^l *as in Proposition 4.5. Then the following equality holds:*

$$P(\mathfrak{z}|(e, \alpha), \lambda) = \begin{cases} \overline{F}_{B^b}(\lambda) \overline{F}_{B^a}(\lambda), & \textit{if } l = 0 \textit{ and } \mathfrak{z} = 0, \\ \overline{F}_{B^b}(\lambda) \overline{F}_{C^1}(\lambda), & \textit{if } l \geq 1 \textit{ and } \mathfrak{z} = 0, \\ \overline{F}_{B^b}(\lambda) \left[F_{C^{\mathfrak{z}}}(\lambda) - \left(F_{C^{\mathfrak{z}}} * F_{\Xi}\right)(\lambda)\right], & \textit{if } l > 1 \textit{ and } \mathfrak{z} \in \{1, \dots, l-1\}, \\ \overline{F}_{B^b}(\lambda) \left[F_{C^l}(\lambda) - F_{A^l}(\lambda)\right], & \textit{if } l \geq 1 \textit{ and } \mathfrak{z} = l, \\ 0, & \textit{otherwise}, \end{cases}$$

where Ξ *is an exponentially distributed random variable with parameter* μ_j^a*, and* $*$ *is the convolution operator.*

4.2. **Laplace method.** Not surprisingly, the distributions of the first-passage time of the generalised birth-death processes A, B, C in Definition 4.3 do not admit closed-form expressions. To compute them, we first determine their Laplace transforms, and invert them numerically. We keep here the notations of Proposition 4.5.

Definition 4.7. Let $f : \mathbb{R}_0^+ \to \mathbb{R}$ be a function absolutely integrable on $[0, \omega]$ for any $\omega > 0$. Its (one-sided) Laplace transform is defined by $\hat{f}(s) := \lim_{\omega \uparrow \infty} \int_0^\omega \mathrm{e}^{-st} f(t) \mathrm{d}t$, for all $s \in \mathbb{C}$ such that the right-hand side converges.

The standard (albeit simplified) inversion formula for the Laplace transform is the Bromwich contour integral, or Mellin inversion [1, Chapter 1]: for an absolutely integrable continuous function f, the identity $f(t) = \frac{1}{2\pi\mathtt{i}}\int_{x-\mathtt{i}\infty}^{x+\mathtt{i}\infty} \mathrm{e}^{ts}\hat{f}(s)\mathrm{d}s$ holds for any $x>0$, and, by symmetry arguments, can be simplified to

$$f(t) = \frac{2\mathrm{e}^{xt}}{\pi}\int_0^\infty \Re\left[\hat{f}(x+\mathtt{i}u)\right]\cos(ut)\mathrm{d}u, \qquad \text{for all } t>0. \tag{4.4}$$

We then apply the Euler algorithm [3, Section 1] that exploits the specific structure of the integrand in (4.4). We now consider the general case of a birth-death process X^b with initial state $b\in\mathbb{N}^+$, and with birth rate $\lambda_n \geq 0$ and death rate $\mu_n > 0$ in state $n\in\mathbb{N}^+$. The following lemma, derived in [12, Equation (14)] following Abate-Whitt methodology [2, Section 4], expresses the Laplace transforms of the density and cumulative distribution function of τ_{X^b}.

Lemma 4.8. *The equality $\hat{F}_{X^b}(s) = s^{-1}\hat{f}_{X^b}(s)$ holds on $\{s\in\mathbb{C} : \Re(s)>0\}$, and*

$$\hat{f}_{X^b}(s) = \prod_{n=1}^{b}\left[-\frac{1}{\lambda_{n-1}}\underset{k\geq 0}{\Phi}\left(\frac{-\lambda_{k+n-1}\mu_{k+n}}{\lambda_{k+n}+\mu_{k+n}+s}\right)\right], \quad \textit{for all } s\in\mathbb{C} \textit{ such that } \Re(s)>0, \tag{4.5}$$

where $\underset{k\geq 0}{\Phi}\dfrac{a_k}{b_k} := \lim_{k\uparrow\infty} t_0\circ t_1\circ\cdots\circ t_k(0)$ *and* $t_k(u) := \dfrac{a_k}{b_k+u}$ *for* $u\geq 0$.

Proposition 4.9. *Fix $v\in\mathbb{N}^+$, $l\in\mathbb{N}$ and $\kappa,\mu,\theta>0$, and denote the processes $A[v,l,\kappa,\mu,\theta]$, $B[v,\kappa,\mu,\theta]$, $C[v,l,\mu,\theta]$ (as in Definition 4.3) by A, B and C, respectively. In particular, we denote the process $B[j,\kappa,\mu,\theta]$ by B^j for any $j\in\mathbb{N}^+$. Assume that f_A, f_B and f_C are continuous on $\mathbb{R}^+$. Then*

$$\hat{f}_B(s) = \frac{1}{(-\kappa)^v}\prod_{n=1}^{v}\underset{k\geq 0}{\Phi}\left[\frac{-\kappa\mu-\kappa(k+n)\theta}{\kappa+\mu+(k+n)\theta+s}\right], \quad \textit{and} \quad \hat{f}_C(s) = \left(\frac{\mu}{\mu+s}\right)^l\prod_{n=l+1}^{l+v}\frac{\mu+(n-l)\theta}{\mu+(n-l)\theta+s},$$

for $\Re(s)>0$. Furthermore, for $R_j(u) := \frac{1}{j!}\exp\left(-\frac{\kappa}{\theta}\left(1-\mathrm{e}^{-\theta u}\right)\right)\left[\frac{\kappa}{\theta}\left(1-\mathrm{e}^{-\theta u}\right)\right]^j$ for $u\geq 0$ and $j\in\mathbb{N}$, we have

$$f_A(t) = f_C(t)R_0(t) + \int_0^t\sum_{j=1}^\infty f_{B^j}(t-u)f_C(u)R_j(u)\mathrm{d}u. \tag{4.6}$$

Proof. The formulae for $\hat{f}_B$ and $\hat{f}_C$ are derived directly from Lemma 4.8, and we therefore focus on (4.6). Let $\tau_\Delta := \tau_A - \tau_C$. Before time τ_C, the process $(G_u) := (G[\kappa,\mu,\theta,\tau_C]_u)$ can be regarded as an initial empty $M/M/\infty$ queue with arrival rate κ and service rate θ. Let $R_j(u)$ denote the probability of G_u being in state $j\in\mathbb{N}$ when $u<\tau_C$. Then, by [32, p. 160], we have

$$R_j(u) = \overline{\mathbb{P}}\left(G_u = j\middle|\tau_C = u\right) = \frac{1}{j!}\exp\left\{-\frac{\kappa}{\theta}\left(1-\mathrm{e}^{-\theta u}\right)\right\}\left[\frac{\kappa}{\theta}\left(1-\mathrm{e}^{-\theta u}\right)\right]^j.$$

Given $\tau_C = u, G_u = j\in\mathbb{N}^+$, the probability density function of τ_Δ is f_{B^j}. Indeed, in the case when $\tau_C = u$ and $G_u = G_{\tau_C} = j$, the time spent on depleting the agent's order and the orders with higher time priority is u and at that time the volume remaining in the queue is of j unit size. The remaining queue can be described by the process B^j, and the depletion time τ_Δ is thus τ_{B^j} (with density f_{B^j}). And given $\tau_C = u, G_u = 0$, we have $\Delta = 0$ almost surely. Therefore, the mixture density $\delta(\cdot)R_0(u)+\sum_{j=1}^\infty f_{B^j}(\cdot)R_j(u)$, with $\delta(\cdot)$ being the Dirac mass, provides the density of τ_Δ given $\tau_C = u$. Furthermore, the function $\delta(\cdot-u)R_0(u)+\sum_{j=1}^\infty f_{B^j}(\cdot-u)R_j(u)(u)$ is the density of $\tau_A = \tau_\Delta+\tau_C$ given $\tau_C = u$. Consequently, for all $t>0$, we have

$$f_A(t) = f_C(t)R_0(t) + \int_0^t\sum_{j=1}^\infty f_{B^j}(t-u)f_C(u)R_j(u)\mathrm{d}u.$$

□

5. Existence of Optimal Policy

We now illustrate our main result, namely the existence and uniqueness of the value function, and the existence of the optimal policy, as in the following theorem.

Theorem 5.1. *The value function in* (3.5) *exists and is unique, and the* $\mathbb{T}$*-optimal policy in* (3.6) *exists.*

The proof of the theorem relies on several ingredients. In the first place, to make a finite-horizon semi-Markov model sensible, it is essential to have a (almost surely) finite number of decision epochs before maturity. In our setting, this is equivalent to the following lemma.

Lemma 5.2. *For any* $(\mathfrak{e},\lambda)\in\mathcal{E}\times\mathbb{T}$, $\pi\in\Pi$, *the limit* $\lim_{n\uparrow\infty}\mathbb{P}^{\pi}_{(\mathfrak{e},\lambda)}(\mathfrak{N}<n)=1$ *holds for* $\mathfrak{N}$ *as in Definition 3.7.*

Proof. According to [21, Proposition 2.1], it suffices to prove that there exist $\zeta,\upsilon>0$ such that

$$Q(\zeta,\mathcal{E}|(e,\alpha))\leq 1-\upsilon, \tag{5.1}$$

for any $(e,\alpha)\in\mathcal{K}$. According to (4.1) and Lemma 4.4, we can write, for any $\zeta>0$ and $(e,\alpha)\in\mathcal{K}$,

$$Q(\zeta,\mathcal{E}|(e,\alpha))=\mathbb{P}^{\pi}_{(\mathfrak{e},\lambda)}(X_{n+1}\leq\zeta|E_n=e,A_n=\alpha)=\overline{\mathbb{P}}\left(\tau_{B^b}\wedge\tau_{A^l}\leq\zeta\right)=1-\overline{\mathbb{P}}\left(\tau_{B^b}>\zeta\right)\overline{\mathbb{P}}\left(\tau_{A^l}>\zeta\right).$$

By Assumption 2.6(c), the agent never consumes up all the volumes at the best bid price through her market order, so that there is at least one unit size order left at the best bid and ask price after the agent's action. By stochastic ordering for the birth and death processes [22, Section 3], the inequalities

$$\overline{\mathbb{P}}\left(\tau_{B^b}>\zeta\right)\geq\overline{\mathbb{P}}\left(\tau_{B[1,0,\mu^b_j,\theta^b_j]}>\zeta\right)\geq\mathrm{e}^{-\iota\zeta},$$

and

$$\overline{\mathbb{P}}\left(\tau_{A^l}>\zeta\right)\geq\overline{\mathbb{P}}\left(\tau_{C^l}>\zeta\right)\geq\overline{\mathbb{P}}\left(\tau_{C[1,0,\mu^a_j,\theta^a_j]}>\zeta\right)\geq\mathrm{e}^{-\iota\zeta},$$

hold with $\iota:=\max\left\{\mu^{\mathfrak{s}}_j+\theta^{\mathfrak{s}}_j:(\mathfrak{s},j)\in\{a,b\}\times\{+1,-1\}\right\}$, and (5.1) therefore holds for $\zeta>0$ and $\upsilon=\mathrm{e}^{-2\iota\zeta}$. □

Next, let $\mathcal{U}$ denote the Banach space of non-negative valued functions on $\mathcal{E}\times\mathbb{T}$ with finite supremum norm:

$$\mathcal{U}:=\left\{u:\mathcal{E}\times\mathbb{T}\to\mathbb{R}^+_0\;\middle|\;\|u\|:=\sup_{(e,\lambda)\in\mathcal{E}\times\mathbb{T}}|u(e,\lambda)|<\infty\right\},$$

and, for any $\pi\in\Pi$, define the dynamic programming operator $\mathcal{T}^\pi$ acting on $\mathcal{U}$ as

$$\mathcal{T}^\pi u(e,\lambda):=r(e,\pi(e,\lambda))+\sum_{\mathfrak{z}=0}^{\infty}w(e,\pi(e,\lambda),\mathfrak{z})P(\mathfrak{z}|(e,\pi(e,\lambda)),\lambda)+\sum_{\tilde{e}\in\mathcal{E}}\int_0^\lambda u(\tilde{e},\lambda-t)Q\big(\mathrm{d}t,\tilde{e}|(e,\pi(e,\lambda))\big),$$

for any $(e,\lambda)\in\mathcal{E}\times\mathbb{T}$. The following proposition, as proved in Appendix C, summarises the properties of $\mathcal{T}^\pi$.

Proposition 5.3. *For any* $\pi\in\Pi$, $\mathcal{T}^\pi$ *is a monotone contraction on* $\mathcal{U}$ *with codomain* $\mathcal{U}$, *and the identity* $V^\pi=\mathcal{T}^\pi V^\pi$ *holds on* $\mathcal{E}\times\mathbb{T}$.

Together with Proposition 5.3, the Banach Fixed-Point's Theorem [18] guarantees the existence and uniqueness of V^π: for any $\pi\in\Pi$, if there exists $u^\pi\in\mathcal{U}$ satisfying $\mathcal{T}^\pi u^\pi=u^\pi$, then $u^\pi=V^\pi$. Introduce now the iteration operator $\mathcal{A}$ acting on $\mathcal{U}$ as, for any $u\in\mathcal{U}$ and $(e,\lambda)\in\mathcal{E}\times\mathbb{T}$,

$$\mathcal{A}u(e,\lambda):=\sup_{\alpha\in\mathcal{A}(e)}\left\{r(e,\alpha)+\sum_{\mathfrak{z}=0}^{\infty}w(e,\alpha,\mathfrak{z})P(\mathfrak{z}|(e,\alpha),\lambda)+\sum_{\tilde{e}\in\mathcal{E}}\int_0^\lambda u(\tilde{e},\lambda-t)Q\big(\mathrm{d}t,\tilde{e}|(e,\alpha)\big)\right\}, \tag{5.2}$$

which is also a contraction with codomain $\mathcal{U}$. Indeed, $\mathcal{A}(\mathcal{U}) \subset \mathcal{U}$ is immediate since the action space is finite, and the contraction property is inherited from that of $\mathcal{T}^{\pi}$ by [14, Theorem 2]. The Banach Fixed-Point's Theorem for contraction mapping ensures that $\mathcal{A}u = u$ has a unique solution, denoted by u^*. Directly from Proposition 5.3 and [14, Theorem 3], we have $u^* = V^*$. Finally, by [26, Section 1], a $\mathbb{T}$-optimal policy is given by, for any $(e, \lambda) \in \mathcal{E} \times \mathbb{T}$,

$$\pi^*(e,\lambda) := \operatorname*{arg\,max}_{\alpha \in \mathcal{A}(e)} \left\{ r(e,\alpha) + \sum_{\mathfrak{z}=0}^{\infty} w(e,\alpha,\mathfrak{z}) P(\mathfrak{z}|(e,\alpha),\lambda) + \sum_{\tilde{e}\in\mathcal{E}} \int_0^{\lambda} V^*(\tilde{e},\lambda - t) Q(\mathrm{d}t, \tilde{e}|(e,\alpha)) \right\}. \tag{5.3}$$

Theorem 5.1 therefore follows.

6. Empirical studies

Our empirical calculations are based on the 'Level-I' LOBSTER data [8, Section 4] for three large-tick stocks: Microsoft (MSFT), Intel (INTC) and Yahoo (YHOO), that are traded on the Nasdaq platform from 11 April 2016 to 15 April 2016, recording all market order arrivals, limit order arrivals, and cancellations at the best prices between 9.30am and 4pm. In order to avoid the impact from the abnormal trading behaviours shortly after market opening or shortly before market closing, we exclude market activities during the first and the last twenty minutes of each trading day. In the following, we first (Section 6.1) illustrate the estimation methodology of the Poisson parameters in Assumption 2.4, as well as the joint distribution of the best volumes after a price change in Assumption 2.2(b). We then (Section 6.2) give a numerical scheme that approximates the value function. We finally (Section 6.3) visualise the optimal policy for liquidating the stock YHOO under different trading conditions.

6.1. **Parameter estimation.**

6.1.1. *Poisson parameters.* As in Assumption 2.1(a), orders from the general market participants are of unit size. We first compute the average size of the market orders, limit orders and cancellations at the best prices, denoted by S^m, S^l and S^c respectively, and choose the unit size to be S^l. Estimation results are given in Table 2.

	MSFT	INTC	YHOO
S^m	140	188	152
S^l	176	317	209
S^c	163	309	201

TABLE 2. average order size (in shares)

We then estimate the Poisson parameters as follows. From historical data, we formulate the set $\mathfrak{Q}_{+1}$ (resp. $\mathfrak{Q}_{-1}$) as the queueing races happening immediately after a price increase (resp. decrease): if the spread is currently one tick, a queueing race $\mathfrak{q}_{+1} \in \mathfrak{Q}_{+1}$ (resp. $\mathfrak{q}_{-1} \in \mathfrak{Q}_{-1}$) starts when the best bid (resp. ask) price increases (resp. decreases) by one tick after the best ask (resp. bid) queue depletes, and ends whenever either the new best ask or bid queue depletes. By maximum likelihood estimation (see Appendix E), we have

$$\hat{\mu}_j^{\mathfrak{s}} = \frac{N_{\mathfrak{s},j}^m}{D_j}\frac{S^m}{S^l}, \qquad \hat{\kappa}_j^{\mathfrak{s}} = \frac{N_{\mathfrak{s},j}^l}{D_j}, \qquad \hat{\theta}_j^{\mathfrak{s}} = \frac{N_{\mathfrak{s},j}^c}{V_{\mathfrak{s},j}}\frac{S^c}{S^l}, \qquad \text{for } \mathfrak{s} \in \{a,b\} \text{ and } j \in \{+1,-1\}, \tag{6.1}$$

where

- $N^m_{\mathfrak{s},j}$, $N^l_{\mathfrak{s},j}$ and $N^c_{\mathfrak{s},j}$ represent the total number of market orders, limit orders and cancellations at $\mathfrak{s}$ price[7] for the queueing races in set $\mathfrak{Q}_j$;
- D_j represents the sum of the length of the queueing races in $\mathfrak{Q}_j$;
- $V_{\mathfrak{s},j} := \sum_{i=1}^{\#\mathfrak{Q}_j} \int_{\mathfrak{T}_i} \mathrm{Vol}^{\mathfrak{s}}_i(t)\mathrm{d}t$, where $\mathrm{Vol}^{\mathfrak{s}}_i(t)$ (resp. $\mathfrak{T}_i$) denotes the volume in unit size at $\mathfrak{s}$ price at time t (resp. the time interval) of the i-th queuing race in $\mathfrak{Q}_j$.

Table 3 gives the Poisson parameter estimation where the agent's action at each decision epoch has no latency. For the three stocks, we find that:

- the rates of market order arrivals are indifferent to the side of the best price and the price move direction;
- immediately after a price increase (resp. decrease), there is a surge of limit order arrivals and cancellations at the best bid (resp. ask) price;
- from an estimation (of the Poisson parameters) point of view, an increase of the price on the bid (resp. ask) side is symmetric to a decrease of price on the ask (resp. bid) side.

Table 4 gives the Poisson parameter estimation where the agent's action at each decision epoch has a one-millisecond latency[8]. By comparing it with Table 3, we observe that:

- the rates of market order arrivals barely change;
- the rates of limit order arrivals and cancellations see a decrease, especially on the bid side after a price increase and on the ask side after a price decrease;
- the symmetry remains unaffected.

		MSFT			INTC			YHOO		
$\mathfrak{s}$	j	μ	κ	θ	μ	κ	θ	μ	κ	θ
a	+1	0.42	3.08	0.31	0.20	2.45	0.16	0.19	1.97	0.26
b	+1	0.40	5.83	0.50	0.23	3.54	0.20	0.20	3.45	0.31
a	-1	0.44	5.84	0.51	0.24	3.82	0.22	0.19	3.21	0.33
b	-1	0.42	3.07	0.31	0.23	2.22	0.16	0.21	1.93	0.21

Table 3. Poisson parameter estimation with no latency

		MSFT			INTC			YHOO		
$\mathfrak{s}$	j	μ	κ	θ	μ	κ	θ	μ	κ	θ
a	+1	0.42	2.89	0.27	0.20	2.36	0.15	0.19	1.87	0.23
b	+1	0.39	3.31	0.40	0.23	2.36	0.17	0.20	2.08	0.25
a	-1	0.42	3.22	0.41	0.24	2.48	0.18	0.18	1.82	0.27
b	-1	0.42	2.87	0.27	0.23	2.14	0.14	0.21	1.83	0.18

Table 4. Poisson parameter estimation with 1ms latency

[7]By abuse of language, 'at a (resp, b price)' means 'at the best ask (resp. bid) price'.

[8]When estimating the Poisson parameters in this case, market activities at the first one millisecond of each queueing race are excluded, and the queueing races with duration shorter than one millisecond are excluded.

6.1.2. *Volume distribution after a price change.* The volume in unit size is approximated by rounding the division of the volume in shares by S^l up to the nearest integer. Figure 1 compares the volume distribution immediately after a price change and one millisecond later for YHOO. We observe that:

- the volume at the best bid (resp. ask) price is quite thin immediately after a price increase (decrease), but see a dramatic increase one millisecond later;
- the volume at the best ask (resp. bid) price keeps the distribution almost unchanged within the first millisecond of the queueing race starting with a price increase (resp. decrease).

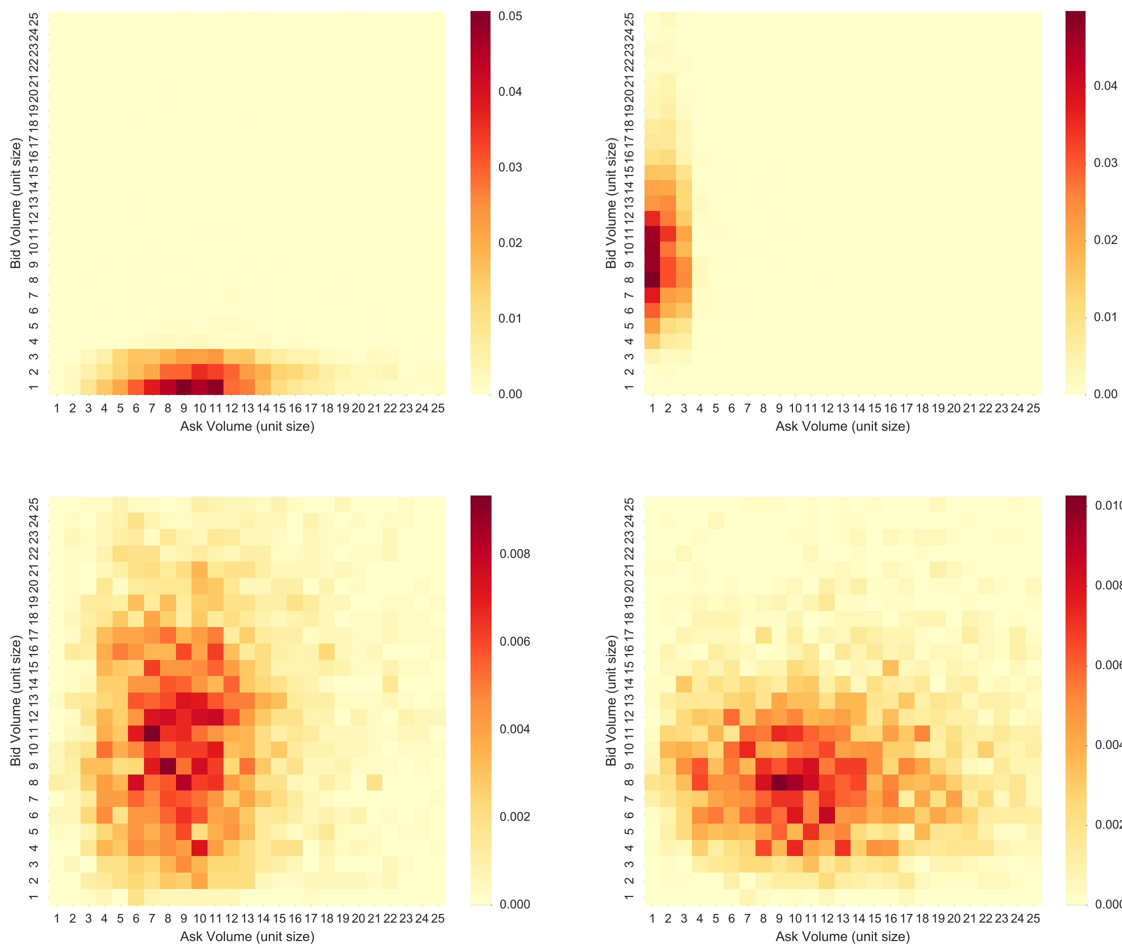

FIGURE 1. YHOO: f_{+1} (left) and f_{-1} (right) with no latency (top) and with 1ms latency (bottom)

6.2. **Numerical scheme.** Dynamic programming techniques usually suffer from the 'curse of dimensionality' [30] to compute the value function through the iteration operator $\mathcal{A}$ in (5.2). The next proposition, as proved in Appendix D, allows us to reduce the dimension of the problem, and hence to accelerate the implementation.

Proposition 6.1. *Given $e := (j, v^b, v^a, p, z, y) \in \mathcal{E}$, $\overline{e} := (j, v^b, v^a, \overline{p}, \overline{z}, y) \in \mathcal{E}$ and $\lambda \in \mathbb{T}$, we have*

$$V^*(e, \lambda) = V^*(\overline{e}, \lambda) + \rho(p - \overline{p})(y + z) + \rho(z - \overline{z})(\overline{p} - j).$$

Besides, the value function V^* is monotone with respect to time to maturity. Indeed, let π^* be $\mathbb{T}$-optimal and construct a policy π^δ, for fixed $\delta \in (0, T)$, as

$$\pi^\delta(e, \lambda) := \begin{cases} \pi^*(e, \lambda + \delta), & \text{if } \lambda \leq T - \delta, \\ (0, 0), & \text{if } \lambda > T - \delta. \end{cases}$$

Definition 3.10 immediately implies that $V^{\pi^\delta}(e, \lambda) \leq V^*(e, \lambda)$ for any $(e, \lambda) \in \mathcal{E} \times \mathbb{T}$ and $V^{\pi^\delta}(e, \lambda) = V^*(e, \lambda - \delta)$ for any $(e, \lambda) \in \mathcal{E} \times [\delta, T]$. The monotonicity in time to maturity therefore follows since δ is arbitrary. As in [24, 27], we can take advantage of the monotonicity of the value function to get a faster convergence rate. The implementation procedure proceeds as follows, for some tolerance level $\mathfrak{tol}$:

Step 1. (initialization): let $n = 0$ and $V_0(e, \lambda) = \rho(p - 1)y + \lambda\rho y/T$ for every $(e, \lambda) \in \mathcal{E} \times \mathbb{T}$;

Step 2. (iteration): choose a random pair $(e_n, \lambda_n) \in \mathcal{E} \times \mathbb{T}$ and compute $\widehat{V}_n := \mathcal{A}V_n(e_n, \lambda_n)$;

Step 3. (correction): with $\widehat{U}_n := \gamma\widehat{V}_n + (1 - \gamma)V_n(e_n, \lambda_n)$ for $\gamma \in (0, 1)$, define the monotonicity projection as:

$$V_{n+1}(e, \lambda) = \begin{cases} \widehat{U}_n, & \text{if } e = e_n, \lambda = \lambda_n, \\ \widehat{U}_n \vee V_n(e, \lambda), & \text{if } e = e_n, \lambda > \lambda_n, \\ \widehat{U}_n \wedge V_n(e, \lambda), & \text{if } e = e_n, \lambda < \lambda_n, \\ V_n(e, \lambda), & \text{if } e \neq e_n; \end{cases}$$

Step 4. (accuracy control): if $\|V_{n+1} - V_n\| \leq \mathfrak{tol}$, end the scheme; otherwise go to Step 2 incrementing n to $n+1$.

6.3. **Optimal policy.** In this section, we provide the results of the optimal policy computed through Formula (5.3) in which the value function is approximated through the numerical scheme in Section 6.2. Figure 2 and 3 illustrate how different market conditions (time to maturity, latency, price move direction, volumes at the best prices) can affect the agent's optimal action (expressed by different colours). We find that posting limit order tends to be prefered when there is more time to maturity, when there is less latency, after a price decrease or when the best ask price has less volume relative to the best bid price.

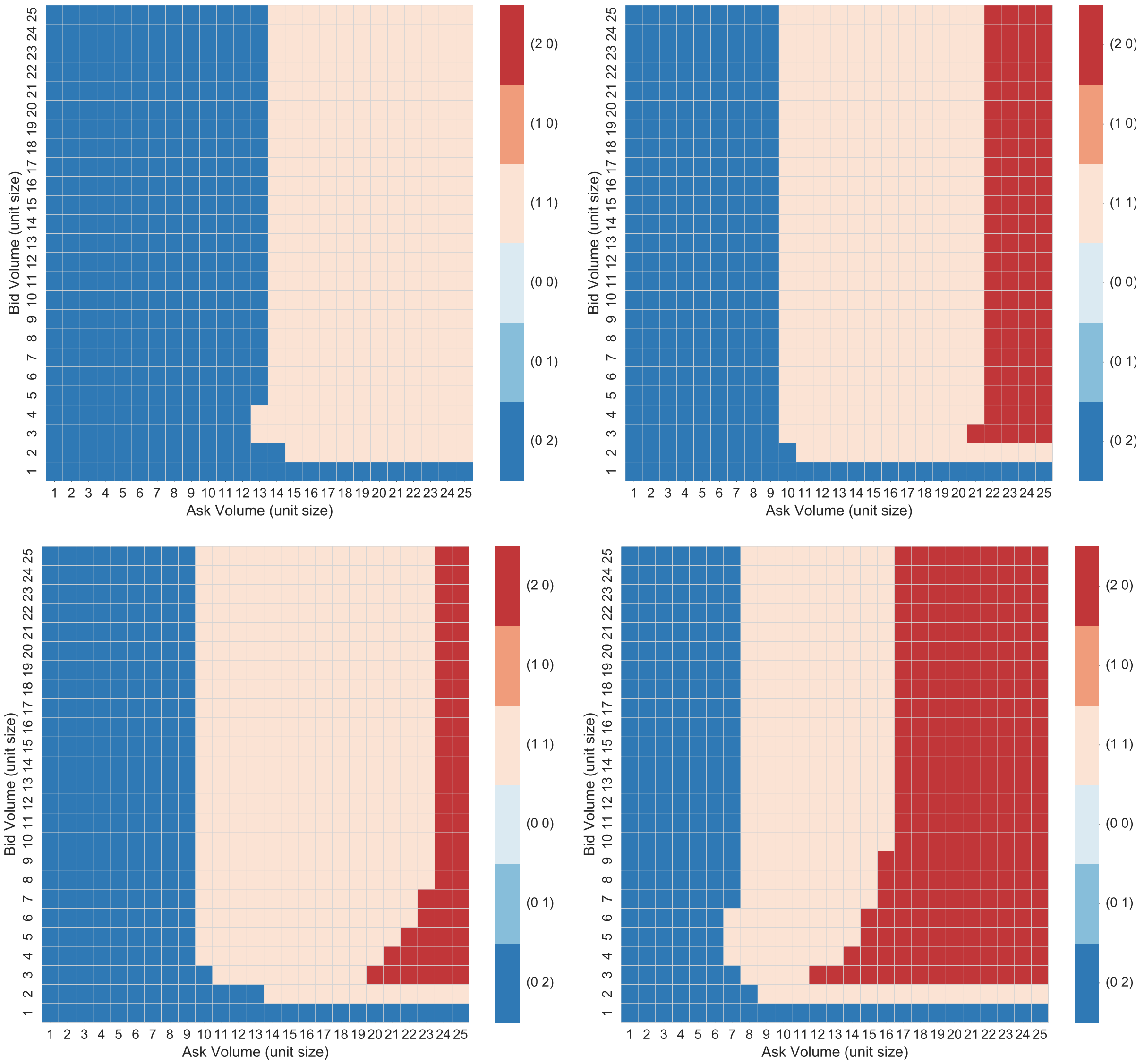

FIGURE 2. Impact of latency (top: no latency; bottom: 1ms latency) and price move direction (left: $j = -1$; right: $j = 1$) when fixing inventory $y = 2$ and time to maturity $\lambda = 10$.

APPENDIX A. PROOF OF PROPOSITION 4.5

In Scenario [S1], the agent posts a limit order at the best ask price ($l \geq 1$), and the best ask queue is depleted before the best bid queue ($\tilde{j} = +1$). Hence,

- the execution time of the best ask queue is less than that of the best bid queue;
- the limit order posted by the agent must get fully executed in the queueing race;
- the duration of the queueing race is the depletion time of the best ask queue.

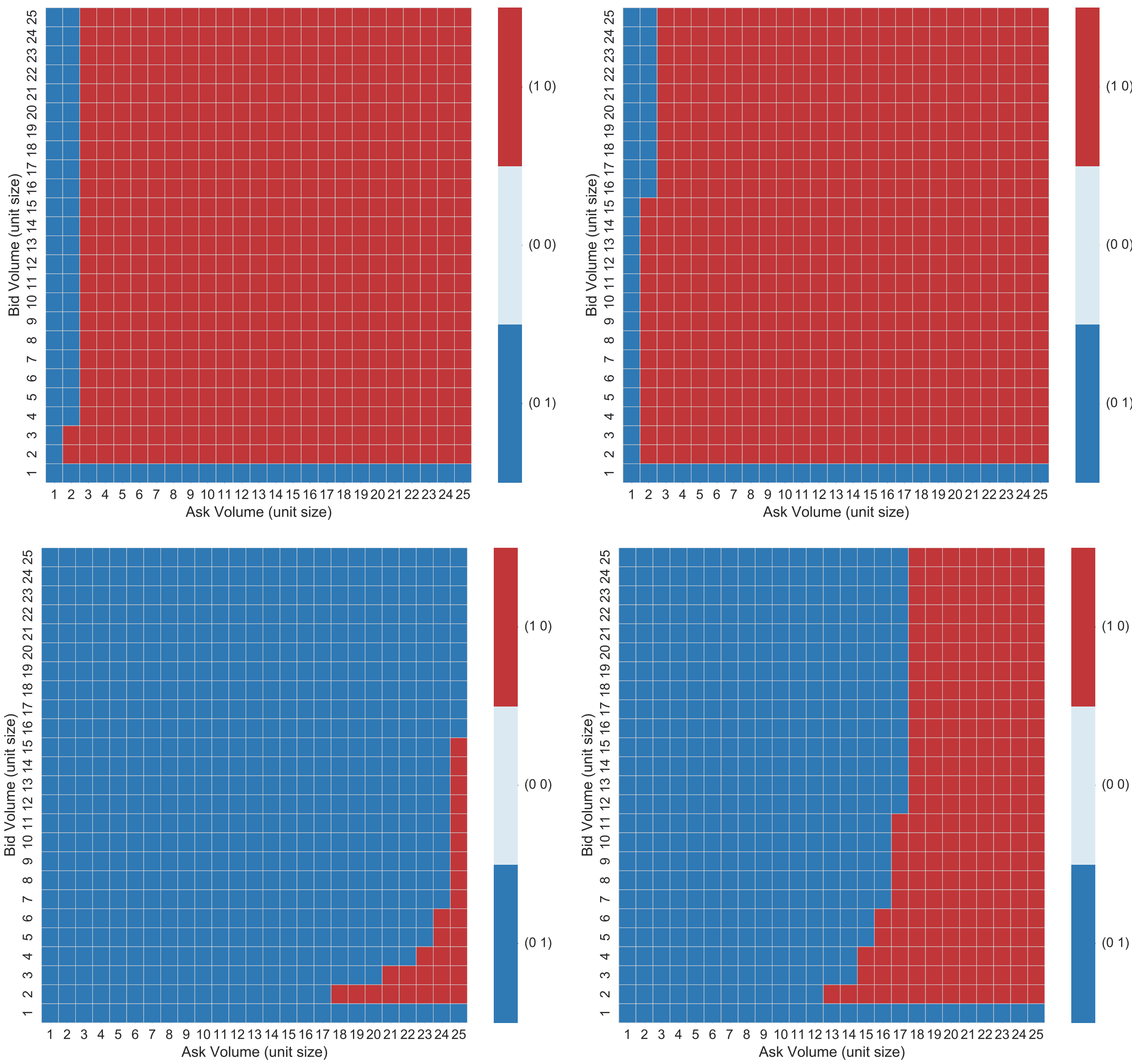

FIGURE 3. Impact of time to maturity (top: $\lambda = 3$; bottom: $\lambda = 10$) and price move direction (left: $j = -1$; right: $j = 1$) when fixing inventory $y = 1$ and 1ms latency.

Therefore, we can write

$$\mathcal{Q}_{j,v,\alpha}\left(t, \tilde{j}, \tilde{z}\right) = \mathbb{P}\left(X_{n+1} \leq t, J_{n+1} = +1 \middle| J_n = j, (V_n^b, V_n^a) = (v^b, v^a), A_n = \alpha\right) \mathbb{P}\left(Z_{n+1} = \tilde{z} \middle| J_{n+1} = +1, L_n = l\right)$$

$$= \overline{\mathbb{P}}\left(\tau_{B^b} > \tau_{A^l}, \tau_{A^l} \leq t\right) \mathbb{1}_{\{\tilde{z}=l\}} = \left\{F_{A^l}(t) - \int_0^t f_{A^l}(u) F_{B^b}(u) \mathrm{d}u\right\} \mathbb{1}_{\{\tilde{z}=l\}}.$$

In Scenario [S2], the agent posts no limit order at the best ask price ($l = 0$). The dynamics of best ask queue can be then described by the process B^a, independent of that of the best bid queue. The proof is similar to that in Scenario [S1].

In Scenario [S3], the agent posts a limit order at the best ask price ($l \geq 0$), and the best bid queue is depleted before the best ask queue ($\tilde{j} = -1$), while the agent's limit order gets no execution ($\tilde{z} = 0$). Hence,

- the execution time of the best bid queue is less than that of one unit size of the agent's limit order together with the limit orders with higher time priority at the best ask price, and is therefore less than that of the entire best ask queue;
- the duration of the queueing race is the depletion time of the best bid queue.

We then have

$$\begin{aligned}\mathcal{Q}_{j,v,\alpha}\left(t,\tilde{j},\tilde{z}\right) &= \mathbb{P}\left(X_{n+1}\leq t, Z_{n+1}=0\middle|J_n=j,(V_n^b,V_n^a)=(v^b,v^a),A_n=\alpha\right)\mathbb{P}\left(J_{n+1}=-1|L_n=l,Z_{n+1}=0\right)\\ &= \overline{\mathbb{P}}\left(\tau_{C^1}>\tau_{B^b},\tau_{B^b}\leq t\right)=F_{B^b}(t)-\int_0^t f_{B^b}(u)F_{C^1}(u)\mathrm{d}u.\end{aligned}$$

In Scenario [S4], the agent posts a limit order of one unit size at the best ask price ($l=1$), the best bid queue is depleted before the best ask queue ($\tilde{j}=-1$) and the agent's limit order gets executed ($\tilde{z}=1$). According to Remark 4.2, we have $\mathcal{Q}_{j,v,\alpha}\left(t,\tilde{j},\{0,1\}\right)=\mathcal{Q}_{j,v,\alpha}\left(t,\tilde{j},0\right)+\mathcal{Q}_{j,v,\alpha}\left(t,\tilde{j},1\right)$, so that

$$\begin{aligned}\mathcal{Q}_{j,v,\alpha}\left(t,-1,1\right) &= \mathbb{P}\left(X_{n+1}\leq t, J_{n+1}=-1\middle|J_n=j,(V_n^b,V_n^a)=(v^b,v^a),A_n=(m,1)\right)-\mathcal{Q}_{j,v,\alpha}\left(t,-1,0\right)\\ &= \overline{\mathbb{P}}\left(\tau_{A^1}>\tau_{B^b},\tau_{B^b}\leq t\right)-\mathcal{Q}_{j,v,\alpha}\left(t,-1,0\right)\\ &= F_{B^b}(t)-\int_0^t f_{B^b}(u)F_{A^1}(u)\mathrm{d}u-\left[F_{B^b}(t)-\int_0^t f_{B^b}(u)F_{C^1}(u)\mathrm{d}u\right]\\ &= \int_0^t f_{B^b}(u)\left[F_{C^1}(u)-F_{A^1}(u)\right]\mathrm{d}u.\end{aligned}$$

In Scenario [S5], the best bid queue is depleted before the best ask queue ($\tilde{j}=-1$), while $\tilde{z}\in\{1,\ldots,l-1\}$ out of $l>1$ unit size of the agent's limit order gets executed when this queueing race terminates. Hence,

- the execution time of the best bid queue lies within the interval $[\tau_{C^{\tilde{z}}},\tau_{C^{\tilde{z}}}+\Delta)$, where Δ is the execution time of one unit size of the agent's limit order when at the top of the queue, which is exponentially distributed with parameter μ_j^a and is independent of $\tau_{C^{\tilde{z}}}$;
- the duration of the queueing race is the depletion time of the best bid queue.

We then have

$$\begin{aligned}&\mathcal{Q}_{j,v,\alpha}\left(t,\tilde{j},\tilde{z}\right)\\ &= \mathbb{P}\left(X_{n+1}\leq t, Z_{n+1}=\tilde{z}\middle|J_n=j,(V_n^b,V_n^a)=(v^b,v^a),A_n=\alpha\right)\mathbb{P}\left(J_{n+1}=-1|L_n^a=l,Z_{n+1}=\tilde{z}\right)\\ &= \overline{\mathbb{P}}\Big(\tau_{C^{\tilde{z}}}\leq\tau_{B^b}<\tau_{C^{\tilde{z}}}+\Delta,\tau_{B^b}\leq t\Big)=\int_0^\infty\int_0^\infty\overline{\mathbb{P}}\left(\tau_{B^b}\in[u,u+\nu),\tau_{B^b}\leq t\right)f_{C^{\tilde{z}}}(u)\overline{\mathbb{P}}(\Delta\in\mathrm{d}\nu)\mathrm{d}u\mathrm{d}\nu\\ &= \int_0^t\int_0^{t-u}\left[F_{B^b}(u+\nu)-F_{B^b}(u)\right]f_{C^{\tilde{z}}}(u)\overline{\mathbb{P}}(\Delta\in\mathrm{d}\nu)\mathrm{d}\nu\mathrm{d}u+\int_0^t\int_{t-u}^\infty\left[F_{B^b}(t)-F_{B^b}(u)\right]f_{C^{\tilde{z}}}(u)\overline{\mathbb{P}}(\Delta\in\mathrm{d}\nu)\mathrm{d}\nu\mathrm{d}u\\ &= \mu_j^a\int_0^t \mathrm{e}^{\mu_j^a u}f_{C^{\tilde{z}}}(u)\left(\int_u^t\mathrm{e}^{-\mu_j^a\epsilon}F_{B^b}(\epsilon)\mathrm{d}\epsilon\right)\mathrm{d}u+\mathrm{e}^{-\mu_j^a t}F_{B^b}(t)\int_0^t\mathrm{e}^{\mu_j^a u}f_{C^{\tilde{z}}}(u)\mathrm{d}u-\int_0^t F_{B^b}(u)f_{C^{\tilde{z}}}(u)\mathrm{d}u\\ &= \mu_j^a\int_0^t F_{B^b}^*(\epsilon)\int_0^\epsilon f_{C^{\tilde{z}}}^*(u)\mathrm{d}u\mathrm{d}\epsilon+\int_0^t\left[F_{B^b}^*(\epsilon)\right]'\int_0^\epsilon f_{C^{\tilde{z}}}^*\mathrm{d}u\mathrm{d}\epsilon\\ &= \int_0^t f_{B^b}^*(\epsilon)\int_0^\epsilon f_{C^{\tilde{z}}}^*(u)\mathrm{d}u\,\mathrm{d}\epsilon\end{aligned}$$

where $f_{C^z}^*(\xi):=\mathrm{e}^{\mu_j^a\xi}f_{C^z}(\xi)$, $f_{B^b}^*(\xi):=\mathrm{e}^{-\mu_j^a\xi}f_{B^b}(\xi)$ and $F_{B^b}^*(\xi):=\mathrm{e}^{-\mu_j^a\xi}F_{B^b}(\xi)$ for any $\xi\geq 0$ and $z\in\mathbb{N}^+$.

Finally, in Scenario [S6], according to Remark 4.2, for $\left(j, v^b, v^a, \alpha\right) \in \mathcal{K}'$ such that $l > 1$, we have

$$\mathcal{Q}_{j,v,\alpha}\left(t, -1, \{0, 1, \dots, l\}\right) = \mathcal{Q}_{j,v,\alpha}\left(t, -1, 0\right) + \sum_{z=1}^{l-1} \mathcal{Q}_{j,v,\alpha}\left(t, -1, z\right) + \mathcal{Q}_{j,v,\alpha}\left(t, -1, l\right),$$

which yields the result by using [S3] and [S5].

Appendix B. Proof of Proposition 4.6

If $l = 0$ and $\mathfrak{z} = 0$, then $P(\mathfrak{z}|(e,\alpha),\lambda) = \overline{\mathbb{P}}\left(\tau_{B^b} \wedge \tau_{B^a} > \lambda\right) = \overline{\mathbb{P}}\left(\tau_{B^b} > \lambda\right)\overline{\mathbb{P}}\left(\tau_{B^a} > \lambda\right) = \overline{F}_{B^b}(\lambda)\overline{F}_{B^a}(\lambda)$.

If $l \geq 1$ and $\mathfrak{z} = 0$, then $P(\mathfrak{z}|(e,\alpha),\lambda) = \overline{\mathbb{P}}\left(\tau_{B^b} \wedge \tau_{A^l} > \lambda, \tau_{C^1} > \lambda\right) = \overline{\mathbb{P}}\left(\tau_{B^b} > \lambda\right)\overline{\mathbb{P}}\left(\tau_{C^1} > \lambda\right) = \overline{F}_{B^b}(\lambda)\overline{F}_{C^1}(\lambda)$.

If $l > 1$ and $\mathfrak{z} \in \{1, \dots, l-1\}$, then

$$\begin{aligned}
P(\mathfrak{z}|(e,\alpha),\lambda) &= \overline{\mathbb{P}}\left(\tau_{B^b} \wedge \tau_{A^l} > \lambda, \tau_{C^\mathfrak{z}} + \Xi > \lambda \geq \tau_{C^\mathfrak{z}}\right) = \overline{\mathbb{P}}\left(\tau_{B^b} > \lambda\right)\overline{\mathbb{P}}\left(\tau_{C^\mathfrak{z}} + \Xi > \lambda \geq \tau_{C^\mathfrak{z}}\right) \\
&= \overline{\mathbb{P}}\left(\tau_{B^b} > \lambda\right)\overline{\mathbb{P}}\left[1 - \overline{\mathbb{P}}\left(\tau_{C^\mathfrak{z}} > \lambda\right) - \overline{\mathbb{P}}\left(\tau_{C^\mathfrak{z}} + \Xi \leq \lambda\right)\right] = \overline{F}_{B^b}(\lambda)\left[F_{C^\mathfrak{z}}(\lambda) - \left(F_{C^\mathfrak{z}} * F_\Xi\right)(\lambda)\right].
\end{aligned}$$

If $l \geq 1$ and $\mathfrak{z} = l$, then

$$\begin{aligned}
P(\mathfrak{z}|(e,\alpha),\lambda) &= \overline{\mathbb{P}}\left(\tau_{B^b} \wedge \tau_{A^l} > \lambda, \tau_{C^l} \leq \lambda\right) = \overline{\mathbb{P}}\left(\tau_{B^b} > \lambda\right)\overline{\mathbb{P}}\left(\tau_{A^l} > \lambda \geq \tau_{C^l}\right) \\
&= \overline{\mathbb{P}}\left(\tau_{B^b} > \lambda\right)\left[1 - \overline{\mathbb{P}}\left(\tau_{C^l} > \lambda\right) - \overline{\mathbb{P}}\left(\tau_{A^l} \leq \lambda\right)\right] = \overline{F}_{B^b}(\lambda)\left[F_{C^l}(\lambda) - F_{A^l}(\lambda)\right].
\end{aligned}$$

According to Remark 3.2, the terminal kernel has zero value in all other scenarios.

Appendix C. Proof of Proposition 5.3

For any $\pi \in \Pi$ and $u \in \mathcal{U}$, we have

$$\|\mathcal{T}^\pi u\| \leq \sup_{(e,\lambda)\in\mathcal{E}\times\mathbb{T}} |r(e,\pi(e,\lambda))| + \sup_{\substack{(e,\lambda)\in\mathcal{E}\times\mathbb{T} \\ \mathfrak{z}\in\{0,\dots,\bar{l}\}}} |w(e,\pi(e,\lambda),\mathfrak{z})| + \sup_{(e,\lambda)\in\mathcal{E}\times\mathbb{T}} \left|\sum_{\tilde{e}\in\mathcal{E}} \int_0^\lambda u(\tilde{e},\lambda - t) Q\big(\mathrm{d}t, \tilde{e}|(e,\pi(e,\lambda))\big)\right|.$$

The first two terms are bounded since the state space $\mathcal{E}$ and the action space $\mathcal{A}$ are finite. Regarding the last term, Lemma 5.2 yields

$$\begin{aligned}
\sup_{(e,\lambda)\in\mathcal{E}\times\mathbb{T}} \left|\sum_{\tilde{e}\in\mathcal{E}} \int_0^\lambda u(\tilde{e},\lambda - t) Q\big(\mathrm{d}t, \tilde{e}|(e,\pi(e,\lambda))\big)\right| &\leq \|u\| \sup_{(e,\lambda)\in\mathcal{E}\times\mathbb{T}} \left|\sum_{\tilde{e}\in\mathcal{E}} \int_0^\lambda Q\big(\mathrm{d}t, \tilde{e}|(e,\pi(e,\lambda))\big)\right| \\
&= \|u\| \sup_{(e,\lambda)\in\mathcal{E}\times\mathbb{T}} Q\big(\lambda, \mathcal{E}|(e,\pi(e,\lambda))\big) \\
&\leq \|u\| \sup_{\lambda\in\mathbb{T}} (1 - \mathrm{e}^{-2\iota\lambda}) = \|u\|(1 - \mathrm{e}^{-2\iota T}) < \infty,
\end{aligned}$$

so that the codomain of $\mathcal{T}^\pi$ is $\mathcal{U}$. The contraction property follows directly from (5.1), since the inequality $\|\mathcal{T}^\pi u - \mathcal{T}^\pi v\| \leq (1 - \mathrm{e}^{-2\iota T})\|u - v\|$ holds for all $u, v \in \mathcal{U}$, and the monotonicity follows from the properties of the semi-Markov kernel. Finally, from (3.7) and Theorem 3.9, we can write, for any $(e,\lambda)\in\mathcal{E}\times\mathbb{T}$,

$$\begin{aligned}
V^\pi(e,\lambda) &= \sum_{n=0}^{\infty}\mathbb{E}^\pi_{(e,\lambda)}\left[r(\mathcal{E}_n,\mathcal{A}_n)\mathbb{1}_{\{\Lambda_n\geq 0\}} + w(\mathcal{E}_n,\mathcal{A}_n,\mathfrak{Z})\mathbb{1}_{\{0\leq\Lambda_n<X_{n+1}\}}\right]\\
&= \mathbb{E}^\pi_{(e,\lambda)}\left[r(\mathcal{E}_0,\mathcal{A}_0)\mathbb{1}_{\{\Lambda_0\geq 0\}} + w(\mathcal{E}_0,\mathcal{A}_0,\mathfrak{Z})\mathbb{1}_{\{0\leq\Lambda_0<X_1\}}\right]\\
&\qquad\qquad + \sum_{n=1}^{\infty}\mathbb{E}^\pi_{(e,\lambda)}\left[r(\mathcal{E}_n,\mathcal{A}_n)\mathbb{1}_{\{\Lambda_n\geq 0\}} + w(\mathcal{E}_n,\mathcal{A}_n,\mathfrak{Z})\mathbb{1}_{\{0\leq\Lambda_n<X_{n+1}\}}\right]\\
&= \mathbb{E}^\pi_{(e,\lambda)}\left[\mathbb{E}^\pi_{(e,\lambda)}\left[r(\mathcal{E}_0,\mathcal{A}_0)\mathbb{1}_{\{\Lambda_0\geq 0\}} + \sum_{\mathfrak{z}=0}^{\infty} w(\mathcal{E}_0,\mathcal{A}_0,\mathfrak{z})\mathbb{1}_{\{0\leq\Lambda_0<X_1,\mathfrak{Z}=\mathfrak{z}\}}\middle| H_0\right]\right]\\
&\qquad\qquad + \sum_{n=1}^{\infty}\mathbb{E}^\pi_{(e,\lambda)}\left[\mathbb{E}^\pi_{(e,\lambda)}\left[r(\mathcal{E}_n,\mathcal{A}_n)\mathbb{1}_{\{\Lambda_n\geq 0\}} + w(\mathcal{E}_n,\mathcal{A}_n,\mathfrak{Z})\mathbb{1}_{\{0\leq\Lambda_n<X_{n+1}\}}\middle| H_1,\right]\right]\\
&= r(e,\pi(e,\lambda)) + \sum_{\mathfrak{z}=0}^{\infty} w(e,\pi(e,\lambda),\mathfrak{z})P(\mathfrak{z}|(e,\pi(e,\lambda)),\lambda)\\
&\qquad + \sum_{n=1}^{\infty}\mathbb{E}^\pi_{(e,\lambda)}\left[\mathbb{E}^\pi_{(\mathcal{E}_1,\Lambda_1)}\left[r(\mathcal{E}_{n-1},\mathcal{A}_{n-1})\mathbb{1}_{\{\Lambda_{n-1}\geq 0\}} + w(\mathcal{E}_{n-1},\mathcal{A}_{n-1},\mathfrak{Z})\mathbb{1}_{\{0\leq\Lambda_{n-1}<X_n\}}\right]\right]\\
&= r(e,\pi(e,\lambda)) + \sum_{\mathfrak{z}=0}^{\infty} w(e,\pi(e,\lambda),\mathfrak{z})P(\mathfrak{z}|(e,\pi(e,\lambda)),\lambda) + \mathbb{E}^\pi_{(e,\lambda)}\left[V^\pi(\mathcal{E}_1,\Lambda_1)\right]\big)\\
&= r(e,\pi(e,\lambda)) + \sum_{\mathfrak{z}=0}^{\infty} w(e,\pi(e,\lambda),\mathfrak{z})P(\mathfrak{z}|(e,\pi(e,\lambda)),\lambda) + \sum_{\tilde{e}\in\mathcal{E}}\int_0^\lambda V^\pi(\tilde{e},\lambda-t)Q\big(\mathrm{d}t,\tilde{e}|(e,\pi(e,\lambda))\big).
\end{aligned}$$

Appendix D. Proof of Proposition 6.1

Let $\mathbf{i} := (0,0,0,1,0,0)$ and $\mathbf{k} := (0,0,0,0,1,0)$, so that $e = \overline{e} + \Delta_p\mathbf{i} + \Delta_z\mathbf{k}$, where $\Delta_p = p - \overline{p}$ and $\Delta_z := z - \overline{z}$. Define further $\hat{e} := \overline{e} + \Delta_z\mathbf{k}$. According to Proposition 5.3 and [14, Theorem 3], we can write

$$V^*(\hat{e},\lambda) = \mathcal{A}V^*(\hat{e},\lambda) = V^*(\overline{e},\lambda) + \rho\Delta_z(\overline{p} - j). \tag{D.1}$$

With the auxiliary function $u(\mathfrak{e},\lambda) := V^*(\mathfrak{e} - \Delta_p\mathbf{i},\lambda) + \rho\Delta_p(y+z)$ for $(\mathfrak{e},\lambda)\in\mathcal{E}\times\mathbb{T}$, simple calculations yield $\mathcal{A}u(\mathfrak{e},\lambda) = u(\mathfrak{e},\lambda)$ for any $(\mathfrak{e},\lambda)\in\mathcal{E}\times\mathbb{T}$, and Theorem 5.1 implies that $V^*(e,\lambda) = V^*(\hat{e},\lambda) + \rho\Delta_p(y+z)$. Combining this with (D.1) finishes the proof.

Appendix E. Maximum Likelihood Estimation for the Poisson Parameters

Fix $\mathfrak{s}\in\{a,b\}$, $j\in\{+1,-1\}$ and denote the Poisson parameters $\mu^{\mathfrak{s}}_j,\kappa^{\mathfrak{s}}_j,\theta^{\mathfrak{s}}_j$ by μ,κ,θ respectively. Introduce the auxiliary parameters $\mu' := \mu S^l/S^m$ and $\theta' := \theta S^l/S^c$. Suppose we observe l_i times of limit order arrivals, m_i times of market order arrivals and c_i times of cancellations on the $\mathfrak{s}$ side in the i-th queueing race, whose starting time is τ_i, duration is d_i and the volume in unit size at $\mathfrak{s}$ price at time t is $\mathrm{Vol}_i(t)$, for $i\in\{1,\ldots,\#\mathfrak{Q}_j\}$.

The likelihood functions are then constructed as:

$$\mathcal{L}\left(\mu' : m_1, \ldots, m_{\#\mathfrak{Q}_j}, d_1, \ldots, d_{\#\mathfrak{Q}_j}\right) := \prod_{i=1}^{\#\mathfrak{Q}_j} \frac{(\mu' d_i)^{m_i}}{m_i!} \mathrm{e}^{-\mu' d_i},$$

$$\mathcal{L}\left(\kappa : l_1, \ldots, l_{\#\mathfrak{Q}_j}, d_1, \ldots, d_{\#\mathfrak{Q}_j}\right) := \prod_{i=1}^{\#\mathfrak{Q}_j} \frac{(\kappa d_i)^{l_i}}{l_i!} \mathrm{e}^{-\kappa d_i},$$

$$\mathcal{L}\left(\theta' : c_1, \ldots, c_{\#\mathfrak{Q}_j}, \Theta(d_1), \ldots, \Theta(d_{\#\mathfrak{Q}_j})\right) := \prod_{i=1}^{\#\mathfrak{Q}_j} \frac{\Theta(d_i)^{c_i}}{c_i!} \mathrm{e}^{-\Theta(d_i)},$$

where $\Theta(d_i) := \theta' \int_{\tau_i}^{\tau_i+d_i} \mathrm{Vol}_i(t)\mathrm{d}t$. Taking logarithms, and cancelling the derivatives yield the optima (6.1) with

$$N_{s,j}^{\varpi} = \sum_{i=1}^{\#\mathfrak{Q}_j} \varpi_i, \quad \text{for } \varpi \in \{m, l, c\} \qquad D_{s,j} = \sum_{i=1}^{\#\mathfrak{Q}_j} d_i, \qquad V_{s,j} = \sum_{i=1}^{\#\mathfrak{Q}_j} \int_{\tau_i}^{\tau_i+d_i} \mathrm{Vol}_i(t)\mathrm{d}t.$$

References

[1] J. Abate, G. L. Choudhury and W. Whitt. An introduction to numerical transform inversion and its application to probability models. Computational Probability, W. Grassman (ed.), Kluwer, Boston: 257-323, 1999.

[2] J. Abate and W. Ward. Computing Laplace transforms for numerical inversion via continued fractions. *INFORMS Journal on Computing*, 11(4): 394-405, 1999.

[3] J. Abate and W. Whitt. Numerical inversion of Laplace transforms of probability distributions. *ORSA Journal on Computing*, 7(1): 36-43, 1995.

[4] A. Alfonsi, A. Fruth and A. Schied. Optimal execution strategies in limit order books with general shape functions. *Quantitative Finance*, 10(2): 143-157, 2010.

[5] R.B. Ash. Real Analysis and Probability. Academic Press, 2014.

[6] M. Avellaneda and S. Stoikov. High-frequency trading in a limit order book. *Quantitative Finance*, 8(3): 217-224, 2008.

[7] J. Bonart and M. D. Gould. Queue imbalance as a one-tick-ahead price predictor in a limit order book. *Market Microstructure and Liquidity*, 2(02): 1650006, 2016.

[8] J. Bonart and M. D. Gould. Latency and liquidity provision in a limit order book. arXiv:1511.04116, 2016.

[9] Á. Cartea, R. F. Donnelly and S. Jaimungal. Enhancing trading strategies with order book signals. SSRN:2668277, 2015.

[10] R. Cont and A. De Larrard. Price dynamics in a Markovian limit order market. *SIAM Journal on Financial Mathematics*, 4(1): 1-25, 2013.

[11] R. Cont, A. Kukanov and S. Stoikov. The price impact of order book events. *Journal of Fin Econometrics*, 12(1): 47-88, 2014.

[12] R. Cont, S. Stoikov and R. Talreja. A stochastic model for order book dynamics. *Operations Research*, 58(3): 549-563, 2010.

[13] K. Dayri and M. Rosenbaum. Large tick assets: implicit spread and optimal tick size. *Market Microstructure and Liquidity*, 1(1), 2015.

[14] E. V. Denardo. Contraction mappings in the theory underlying dynamic programming. *Siam Review*, 9(2): 165-177, 1967.

[15] P. Fodra and H. Pham. High frequency trading and asymptotics for small risk aversion in a Markov renewal model. *SIAM Journal on Financial Mathematics*, 6(1): 656-684, 2015.

[16] P. Fodra and H. Pham. Semi-Markov model for market microstructure. *Applied Mathematical Finance*, 22(3): 261-295, 2015.

[17] M. Gould, M. Porter, S. Williams, M. McDonald, D. Fenn and S. Howison. Limit order books. *Quantitative Finance*, 13(11): 1709-1742, 2013.

[18] A. Granas and J. Dugundji. Fixed point theory. Springer Science & Business Media, 2013.

[19] F. Guilbaud and H. Pham. Optimal high-frequency trading with limit and market orders. *Quant Finance*, 13(1): 79-94, 2013.

[20] W. Huang, C.-A. Lehalle and M. Rosenbaum. Simulating and analyzing order book data: The queue-reactive model. *Journal of the American Statistical Association*, 110(509): 107-122, 2015.

[21] Y. Huang and X. Guo. Finite horizon semi-Markov decision processes with application to maintenance systems. *European Journal of Operational Research*, 212(1): 131-140, 2011.

[22] A. Irle. Stochastic ordering for continuous-time processes. *Journal of Applied Probability*, 40(2): 361-375, 2003.

[23] P. K. Jain. Financial market design and the equity premium: Electronic versus floor trading. *The Journal of Finance*, 60(6): 2955-2985, 2005.

[24] D. R. Jiang and W. B. Powell. An approximate dynamic programming algorithm for monotone value functions. *Operations Research*, 63(6): 1489-1511, 2015.

[25] C. Maglaras, C. C. Moallemi and H. Zheng. Optimal execution in a limit order book and an associated microstructure market impact model. ssrn:2610808, 2015.

[26] J. W. Mamer. Successive approximations for finite horizon, semi-Markov decision processes with application to asset liquidation. *Operations Research*, 34(4): 638-644, 1986.

[27] J. M. Nascimento and W. B. Powell. An optimal approximate dynamic programming algorithm for the lagged asset acquisition problem. *Mathematics of Operations Research*, 34(1): 210-237, 2009.

[28] Y. Nevmyvaka, Y. Feng and M. Kearns. Reinforcement learning for optimized trade execution. Proceedings of the 23rd international conference on Machine learning, 673-680, 2006.

[29] A. A. Obizhaeva and J. Wang. Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets*, 16(1): 1-32, 2013.

[30] W. B. Powell. Approximate Dynamic Programming: Solving the curses of dimensionality, 703. John Wiley & Sons, 2007.

[31] E. Smith, J. D. Farmer, L. s. Gillemot, S. Krishnamurthy. Statistical theory of the continuous double auction. *Quantitative finance*, 3(6): 481-514, 2003.

[32] L. Takács. Introduction to the Theory of Queues. Oxford University press, Chapman & Hall, 1959.

[33] H. C. Tijms. A first course in stochastic models. John Wiley and Sons, 2003.

[34] L. Zhao. A model of limit-order book dynamics and a consistent estimation procedure. PhD thesis, Carnegie Mellon, 2010.

Department of Mathematics, Imperial College London
E-mail address: `a.jacquier@imperial.ac.uk`

Department of Mathematics, Imperial College London
E-mail address: `hao.liu112@imperial.ac.uk`